Novel selection processes

Description

The present invention relates to processes for preparing transformed plant cells or organisms by transforming a population of plant cells which comprises at least one marker protein having a direct or indirect toxic effect for said population, with at least one nucleic acid sequence to be inserted in combination with at least one compound, preferably a DNA construct, capable of reducing the expression, amount, activity and/or function of the marker protein, with the transformed plant cells having a growth advantage over nontransformed cells, due to the action of said compound.

Genetic material is successfully introduced usually only into a very limited number of target cells of a population. This necessitates the distinction and isolation of successfully transformed from nontransformed cells, a process which is referred to as selection. Traditionally, the selection is carried out by way of a "positive" selection, wherein the transformed cell is enabled to grow and to survive, whereas the untransformed cell is inhibited in its growth or destroyed (McCormick et al. (1986) Plant Cell Reports 5:81-84). A positive selection of this kind is usually implemented by genes which code for a resistance to a biocide (e.g. a herbicide such as phosphinothricin, glyphosate or bromoxynil, a metabolism inhibitor such as 2-deoxyglucose 6-phosphate (WO 98/45456) or an antibiotic such as tetracycline, ampicillin, kanamycin, G 418, neomycin, bleomycin or hygromycin). Such genes are also referred to as positive selection markers. The positive selection marker is coupled (physically or by means of cotransformation) to the nucleic acid sequence to be introduced into the cell genome and is then introduced into the cell. Subsequently, the cells are cultured on a medium under the appropriate selection pressure (for example in the presence of an appropriate antibiotic or herbicide), whereby the transformed cells, owing to the required resistance to said selection pressure, have a growth/survival advantage and can thus be selected. Positive selection markers which may be mentioned by way of example are:

- phosphinothricin acetyltransferases (PAT) (also: Bialophos® resistance; bar) acetylate the free amino group of the glutamine synthase inhibitor phosphinothricin (PPT) and thus achieve a detoxification (de Block et al. (1987) EMBO J

6:2513-2518; Vickers JE et al. (1996) Plant Mol Biol Reporter 14:363–368; Thompson CJ et al. (1987) EMBO J 6:2519–2523).

- 5-enolpyruvylshikimate 3-phosphate synthases (EPSPS) impart a resistance to the unselective herbicide Glyphosat® (N-(phosphonomethyl)glycine; Steinrucken HC et al. (1980) Biochem Biophys Res Commun 94:1207–1212; Levin JG and Sprinson DB (1964) J Biol Chem 239:1142–1150; Cole DJ (1985) Mode of action of glyphosate; A literature analysis, p. 48–74. In: Grossbard E and Atkinson D (eds.) The herbicide glyphosate. Buttersworths, Boston.). Glyphosate-tolerant EPSPS variants for use as selection markers have been described (Padgette SR et al. (1996). New weed control opportunities: development of soybeans with a Roundup Ready™ gene. In: Herbicide Resistant Crops (Duke SO, ed.), pp. 53-84. CRC Press, Boca Raton, FL; Saroha MK and Malik VS (1998) J Plant Biochemistry and Biotechnology 7:65–72; Padgette SR et al.(1995) Crop Science 35(5):1451–1461; US 5,510,471; US 5,776,760; US 5,864,425; US 5,633,435; US 5,627,061; US 5,463,175; EP-A 0 218 571).

- neomycin phosphotransferases constantly impart a resistance to aminoglycoside antibiotics such as neomycin, G418, hygromycin, paromomycin or kanamycin by reducing the inhibiting action thereof by means of a phosphorylation reaction (Beck et al. (1982) Gene 19:327-336).

- 2-deoxyglucose 6-phosphate phosphatases impart a resistance to 2-deoxyglucose (EP-A 0 807 836; Randez-Gil et al. (1995) Yeast 11:1233-1240; Sanz et al. (1994) Yeast 10:1195-1202).

- acetolactate synthases impart a resistance to imidazolinone/sulfonylurea herbicides (e.g. imazzamox, imazapyr, imazaquin, imazethapyr, amidosulforon, azimsulfuron, chlorimuron ethyl, chlorsulfuron; Sathasivan K et al. (1990) Nucleic Acids Res 18(8):2188).

In addition, resistance genes to the antibiotics hygromycin (hygromycin phosphotransferases), chloramphenicol (chloramphenicol acetyltransferase), tetracycline, streptomycin, zeocine and ampicillin (ß-lactamase gene; Datta N, Richmond MH.(1966) Biochem J 98(1):204-9) have been described.

Genes such as isopentenyl transferase (ipt) from Agrobacterium tumefaciens (strain:PO22) (GenBank Acc. No.: AB025109) may likewise be used as selection markers. The ipt gene is a key enzyme of cytokine biosynthesis. Its overexpression facilitates the re-

generation of plants (e.g. selection on cytokine-free medium) (Ebinuma H et al. (2000) Proc Natl Acad Sci USA 94:2117-2121; Ebinuma H *et al.* (2000) Selection of Marker-free transgenic plants using the oncogenes (*ipt, rol* A, B, C) of Agrobacterium as selectable markers, In Molecular Biology of Woody Plants. Kluwer Academic Publishers). The disadvantages here are, firstly, the fact that the selection disadvantage is based on usually subtle differences in cell proliferation and, secondly, the fact that the plant acquires unwanted properties (gall tumor formation) due to transformation with an oncogene.

EP-A 0 601 092 describes various other positive selection markers. Examples which may be mentioned are: β-glucuronidase (in connection with, for example, cytokinine glucuronide), mannose 6-phosphate isomerase (in connection with mannose), UDP-galactose 4-epimerase (in connection with galactose, for example).

Negative selection markers are used for selecting organisms in which marker sequences have been successfully deleted (Koprek T et al. (1999) Plant J 19(6):719-726). In the presence of a negative selection marker, the corresponding cell is destroyed or experiences a growth disadvantage. Negative selection involves, for example, the negative selection marker introduced into the plant converting a compound which otherwise has no action disadvantageous to the plant into a compound with a disadvantageous (i.e. toxic) action. Examples of negative selection markers include: thymidine kinase (TK), for example of Herpes simplex virus (Wigler et al. (1977) Cell 11:223), cellular adenine phosphoribosyl transferase (APRT) (Wigler et al. (1979) Proc Natl Acad Sci USA 76:1373), hypoxanthine phosphoribosyl transferase (HPRT) (Jolly et al. (1983) Proc Natl Acad Sci USA 80:477), diphtheria toxin A fragment (DT-A), the bacterial xanthine-guanine phosphoribosyl transferase (gpt; Besnard et al. (1987) Mol. Cell. Biol. 7:4139; Mzoz and Moolten (1993) Human Gene Therapy 4:589-595), the codA gene product coding for a cytosine deaminase (Gleave AP et al. (1999) Plant Mol Biol. 40(2):223-35; Perera RJ et al. (1993) Plant Mol Biol 23(4): 793-799; Stougaard J; (1993) Plant J 3:755-761; EP-A1 595 873), the cytochrome P450 gene (Koprek et al. (1999) Plant J 16:719-726), genes coding for a haloalkane dehalogenase (Naested H (1999) Plant J 18:571-576), the iaaH gene (Sundaresan V et al. (1995) Genes & Development 9:1797-1810) or the tms2 gene (Fedoroff NV & Smith DL (1993) Plant J 3: 273-289). The negative selection markers are usually employed in combination with "prodrugs" or "pro-toxins", compounds which are converted into toxins by the activity of the selection marker.

5-Methylthioribose (MTR) kinase is an enzyme whose enzymic activity in plants, bacteria and protozoa, but not in mammals, has been described. The enzyme may convert an MTR analog (5-(triromethyl)thioribose) as a "subversive substrate" of the methionine salvage pathway via an unstable intermediate to give the toxic compound carbothionyl difluoride.

Said selection systems have various disadvantages. The introduced selection marker (e.g. resistance to antibiotics) is justified only during transformation and selection but is later a usually unnecessary and often also undesired protein product. This may be disadvantageous for reasons of consumer acceptance and/or approval as a food and/or feed product. Another disadvantage in this connection is the fact that the selection marker used for selection is usually genetically coupled to the nucleic acid sequence to be inserted into the genome and cannot be decoupled by segregation during propagation or crossing. Usually, deletion of the marker sequence is required, making additional steps necessary. In addition, biotechnological studies require in numerous cases multiple transformation with various gene constructs. Here, each transformation step requires a new selection marker unless the previously used marker is to be laboriously deleted first. This, however, necessitates a broad palette of well-functioning selection markers which are not available for most plant organisms.

Consequently, it was the object of the invention to provide novel selection processes for selecting transformed plant cells and organisms, which, if possible, no longer have the disadvantages of the available systems. This object is achieved by the present invention.

The invention firstly relates to a process for preparing transformed plant cells or organisms, which process comprises the following steps:

a) transforming a population of plant cells, with the cells of said population containing at least one marker protein capable of causing directly or indirectly a toxic effect for said population, with at least one nucleic acid sequence to be inserted in combination with at least one compound capable of reducing the expression, amount, activity and/or function of at least one marker protein, and

b) selecting transformed plant cells whose genome contains said nucleic acid sequence and which have a growth advan-

tage over nontransformed cells, due to the action of said compound, from said population of plant cells, the selection being carried out under conditions under which the marker protein can exert its toxic effect on the nontransformed cells.

In a preferred embodiment, the marker protein is a protein capable of converting directly or indirectly a substance X which is nontoxic for said population of plant cells into a substance Y which is toxic for said population. In this case, the process of the invention preferably comprises the following steps:

a) transforming the population of plant cells with at least one nucleic acid sequence to be inserted in combination with at least one compound capable of reducing the expression, amount, activity and/or function of at least one marker protein, and

b) treating said population of plant cells with the substance X at a concentration which causes a toxic effect for nontransformed cells, due to the conversion by the marker protein, and

c) selecting transformed plant cells whose genome contains said inserted nucleic acid sequence and which have a growth advantage over nontransformed cells, due to the action of said compound, from said population of plant cells, the selection being carried out under conditions under which the marker protein can exert its toxic effect on the nontransformed cells.

The nontoxic substance X is preferably a substance which does not naturally occur in plant cells or organisms or occurs naturally therein only at a concentration which can essentially not cause any toxic effect. In the scope of the process of the invention, preference is given to applying the nontoxic substance X exogenously, for example via the medium or the growth substrate.

The term "compound capable of reducing the expression, amount, activity and/or function of at least one marker protein" is to be understood broadly and generally means any compounds which cause, directly or indirectly, alone or in cooperation with other factors, a reduction in the amount of protein, amount of RNA, gene activity, protein activity or protein function of at least one marker protein. Said compounds are also referred to under the generic term "anti-marker protein" compounds. The term "anti-marker

protein" compound includes in particular, but is not limited to, the nucleic acid sequences, ribonucleic acid sequences, double-stranded ribonucleic acid sequences, antisense ribonucleic acid sequences, expression cassettes, peptides, proteins or other factors used in the preferred embodiments within the scope of the process of the invention.

In a preferred embodiment, "anti-marker protein" compound means a DNA construct comprising

a) at least one expression cassette suitable for expressing a ribonucleic acid sequence and/or, if appropriate, a protein, said nucleic acid sequence and/or protein being capable of reducing the expression, amount, activity and/or function of the marker protein, or

b) at least one sequence which causes a partial or complete deletion or inversion of the sequence coding for said marker protein and thus enables the expression, amount, activity and/or function of the marker protein to be reduced, and also, if appropriate, further functional elements which facilitate and/or promote said deletion or inversion, or

c) at least one sequence which causes an insertion into the sequence coding for said marker protein and thus enables the expression, amount, activity and/or function of the marker protein to be reduced, and also, if appropriate, further functional elements which facilitate and/or promote said insertion.

The process of the invention stops the negative-selective action of the marker protein. To this extent, an "anti-marker protein" compound acts directly (e.g. via inactivation by means of insertion into the gene coding for the marker protein) or indirectly (e.g. by means of the ribonucleic acid sequence expressed via the expression cassette and/or, where appropriate, of the protein translated therefrom) as a positive selection marker. Hence, the selection system of the invention is to be referred to as a "reverse selection system", since it "reverts" the negative-selective action of the marker protein.

The process of the invention means a drastic broadening of the repertoire of positive selection processes for selecting transformed plant cells.

Another advantage is the fact that in a particular, preferred embodiment (e.g. via the action of a double-stranded or antisense RNA), it is possible to implement the selection effect without expressing a foreign protein (see below).

It is also advantageous that the marker protein used indirectly for selection (e.g. the negative selection marker) is not coupled genetically to the nucleic acid sequence to be inserted into the genome. In contrast to the otherwise customary selection processes, the marker protein, if it is a transgene, may be removed by simple segregation in the course of subsequent propagation or crossing.

"Plant cell" means within the scope of the present invention any type of cell which has been derived from a plant organism or is present therein. In this context, the term includes by way of example protoplasts, callus or cell cultures, microspores, pollen, cells in the form of tissues such as leaves, meristem, flowers, embryos, roots, etc. Included are, in particular, all of those cells and cell populations which are suitable as target tissues for a transformation.

In this context, "plant organism" comprises any organism capable of photosynthesis and also the cells, tissues, parts or propagation material (such as seeds or fruits) derived therefrom. Included within the scope of the invention are all genera and species of higher and lower plants of the plant kingdom. Preference is given to annual, perennial, monocotyledonous and dicotyledonous plants and also gymnosperms.

"Plant" means within the scope of the invention all genera and species of higher and lower plants of the plant kingdom. The term includes the mature plants, seed, shoots and seedlings, and also parts, propagation material (for example tubers, seeds or fruits), plant organs, tissues, protoplasts, callus and other cultures, for example cell cultures, derived therefrom, and also any other types of groupings of plant cells to give functional or structural units. Mature plants means plants at any developmental stage beyond that of the seedling. Seedling means a young immature plant at an early developmental stage. "Plant" comprises all annual and perennial monocotyledonous and dicotyledonous plants and includes by way of example but not by limitation those of the genera Cucurbita, Rosa, Vitis, Juglans, Fragaria, Lotus, Medicago, Onobrychis, Trifolium, Trigonella, Vigna, Citrus, Linum, Geranium, Manihot, Daucus, Arabidopsis, Brassica, Raphanus, Sinapis, Atropa, Capsicum, Datura, Hyoscyamus, Lycopersicon,

Nicotiana, Solarium, Petunia, Digitalis, Majorana, Cichorium, Helianthus, Lactuca, Bromus, Asparagus, Antirrhinum, Heterocallis, Nemesis, Pelargonium, Panieum, Pennisetum, Ranunculus, Senecio, Salpiglossis, Cucumis, Browaalia, Glycine, Pisum, Phaseolus, Lolium, Oryza, Zea, Avena, Hordeum, Secale, Triticum, Sorghum, Picea and Populus.

Preference is given to plants of the following plant families: Amaranthaceae, Asteraceae, Brassicaceae, Carophyllaceae, Chenopodiaceae, Compositae, Cruciferae, Cucurbitaceae, Labiatae, Leguminosae, Papilionoideae, Liliaceae, Linaceae, Malvaceae, Rosaceae, Rubiaceae, Saxifragaceae, Scrophulariaceae, Solanacea, Sterculiaceae, Tetragoniacea, Theaceae, Umbelliferae.

Preferred monocotyledonous plants are selected in particular from the monocotyledonous crop plants such as, for example, those in the family of Gramineae such as alfalfa, rice, corn, wheat or other cereal species such as barley, millet, rye, triticale or oats and also from sugar cane and all grass species.

Preferred dicotyledonous plants are selected in particular from the dicotyledonous crop plants such as, for example,

- Asteraceae, such as sunflower, tagetes or calendula and others,

- Compositae, in particular the genus Lactuca, very especially the species sativa (lettuce) and others,

- Cruciferae, especially the genus Brassica, very especially the species napus (oilseed rape), campestris (beet), oleracea cv Tastie (cabbage), oleracea cv Snowball Y (cauliflower) and oleracea cv Emperor (broccoli) and other cabbage species; and the genus Arabidopsis, very especially the species thaliana, and cress or canola and others,

- Cucurbitaceae, such as melon, pumpkin/squash or zucchini and others,

- Leguminosae, especially the genus Glycine, very especially the species max (soybean) and alfalfa, pea, bean plant or peanut, and others

- Rubiaceae, preferably the subclass Lamiidae, such as, for example, Coffea arabica or Coffea liberica (coffee bush) and others,

- Solanaceae, in particular the genus Lycopersicon, very especially the species esculentum (tomato), the genus Solanum, very especially the species tuberosum (potato) and melongena (eggplant), and the genus Capsicum, very especially the species annuum (pepper) and tobacco and others,

- Sterculiaceae, preferably the subclass Dilleniidae, such as, for example, Theobroma cacao (cacao tree) and others,

- Theaceae, preferably the subclass Dilleniidae, such as, for example, Camellia sinensis or Thea sinensis (tea shrub) and others,

- Umbelliferae, especially the genus Daucus (very especially the species carota (carrot)) and Apium (very especially the species graveolens dulce (celery)) and others,

and linseed, cotton, hemp, flax, cucumber, spinach, carrot, sugar beet and the various tree, nut and grapevine species, in particular banana and kiwi.

Plant organisms for the purposes of the invention are furthermore other photosynthetically active capable organisms such as, for example, algae, cyanobacteria and mosses. Preferred algae are green algae such as, for example, algae of the genus Haematococcus, Phaedactylum tricornatum, Volvox or Dunaliella. Particular preference is given to Synechocystis.

Particular preference is given to the group of plants, consisting of wheat, oats, millet, barley, rye, corn, rice, buckwheat, sorghum, triticale, spelt, linseed, sugar cane, oilseed rape, cress, Arabidopsis, cabbage species, soybean, alfalfa, pea, bean plants, peanut, potato, tobacco, tomato, eggplant, paprika, sunflower, tagetes, lettuce, calendula, melon, pumpkin and zucchini.

Most preference is given to

a) plants suitable for producing oil, such as, for example, oilseed rape, sunflower, sesame, safflower (Carthamus tinctorius), olive tree, soybean, corn, peanut, ricinus, oil palm, wheat, cacao tree or various nut species such as, for example, walnut, coconut or almond. Among these, particular preference is in turn given to dicotyledonous plants, in particular oilseed rape, soybean and sunflower.

b) plants suitable for producing starch, such as corn, wheat or potato, for example.

c) plants which are utilized as food and/or feedstuff and/or as useful plants and in which a resistance to pathogens would be advantageous, such as barley, rye, rice, potato, cotton, flax or linseed, for example.

d) plants which may be suitable for producing fine chemicals such as, for example, vitamins and/or carotenoids, such as oilseed rape, for example.

"Population of plant cells" means any group of plant cells, which may be subjected within the scope of the present invention to a transformation and from which transgenic plant cells transformed by the process of the invention may be obtained and isolated. In this context, said population may also be, for example, a plant tissue, organ or a cell culture, etc. Said population may comprise by way of example but not by limitation an isolated zygote, an isolated immature embryo, embryogenic callus, plant or else various flower tissues (both in vitro and in vivo).

"Genome" means the entirety of genetic information of a plant cell and comprises both genetic information of the nucleus and that of the plastids (e.g. chloroplasts) and mitochondria. However, genome preferably means the genetic information of the nucleus (for example of the nuclear chromosomes).

"Selection" means identifying and/or isolating successfully transformed plant cells from a population of nontransformed cells by using the process of the invention. This does not necessarily require that the selection be carried out directly with the transformed cells immediately after transformation. It is also possible to carry out the selection only at a later time, even with a later generation of the plant organisms (or cells, tissues, organs or propagation material derived therefrom) resulting from the transformation. Thus it is possible, for example, to transform Arabidopsis plants directly using, for example, the vacuum infiltration method (Clough S & Bent A (1998) Plant J 16(6):735-43; Bechtold N et al. (1993) CR Acad Sci Paris 1144(2):204-212), which subsequently produce transgenic seeds which may then be subjected to selection.

The fact that the nucleic acid sequence to be inserted is transformed "in combination with" the "anti-marker protein" compound (e.g. a DNA construct) is to be understood broadly and means that

at least one nucleic acid sequence to be inserted and at least one "anti-marker protein" compound are functionally coupled to one another so that the presence of the "anti-marker protein" compound in the plant cell, and of the selection advantage related thereto, indicates the parallel presence of the inserted nucleic acid sequence as likely. The nucleic acid sequence to be inserted and the "anti-marker protein" compound (e.g. a DNA construct) here may be, preferably but not necessarily, part of a single nucleic acid construct (e.g. a transformation construct or transformation vector), i.e. be present physicochemically coupled via a covalent bond. However, they may also be jointly introduced separately, for example in the course of a cotransformation, and exert their function within the scope of the process of the invention also in this way. In the case of the "anti-marker protein compound" acting via expressing an RNA (e.g. an antisense RNA or double-stranded RNA) or being such an RNA, "in combination" may also include those embodiments in which said RNA and the RNA expressed by the nucleic acid sequence inserted into the genome form an RNA strand.

"Nontoxic substance X" generally means substances which, compared to their reaction product Y, under otherwise identical conditions, have a reduced, preferably an essentially lacking biological activity, preferably toxicity. In this context, the toxicity of substance Y is at least twice as high as that of substance X, preferably at least five times as high, particularly preferably at least ten times as high, very particularly preferably at least twenty times as high, most preferably at least one hundred times as high. "Identical conditions" here means that all conditions are kept the same, apart from the different substances X and Y. Accordingly, identical molar concentrations of X and Y are used, with the medium, temperature, type of organism and density of organism, etc. being the same. The substance X may be converted to the substance Y in various ways, for example by hydrolysis, deamination, hydrolysis, dephosphorylation, phosphorylation, oxidation or any other type of activation, metabolization or conversion. The substance X may be, by way of example but not by limitation, the inactive precursor or derivative of a plant growth regulator or herbicide.

"Toxicity" or "toxic effect" means a measurable, negative influence on the physiology of the plant or of the plant cell and may comprise here symptoms such as, for example, but not limited thereto, a reduced or disrupted growth, a reduced or disrupted rate of photosynthesis, a reduced or disrupted cell division, a reduced or disrupted regeneration of a complete plant from cell culture or callus, etc.

The plant cells successfully transformed by means of the process of the invention may, to put it differently, have a growth advantage or selection advantage over the nontransformed cells of the same starting population under the influence of the substance "X". Growth or selection advantage is to be understood here broadly and means, for example, the fact that said transformed plant cells are capable of forming shoots and/or can be regenerated to give complete plants, whereas the nontransformed cells can do this only with a marked delay, if at all.

The term of "marker protein" is to be understood broadly and generally means all of those proteins which are capable of

i) exerting per se a toxic effect on the plant or plant cell, or

ii) converting directly or indirectly a nontoxic substance X into a substance Y which is toxic for the plant or plant cell.

In this context, the marker protein may be a plant-intrinsic, endogenous gene or else a transgene from a different organism. Preferably, the marker protein itself has no essential function for the organism including the marker protein. If the marker protein per se exerts a toxic effect, then it will preferably be expressed, for example, under an inducible promoter rather than constitutively.

Preferably, however, the marker protein converts directly or indirectly a nontoxic substance X into a substance Y which is toxic for the plant or plant cell. Particularly preferred marker proteins are the "negative selection markers" as are used, for example, in the course of targeted deletions from the genome.

Examples of marker proteins which may be mentioned but which are not limiting are:

(a) cytosine deaminases (CodA or CDase), with preference being given to using as the nontoxic substance X substances such as 5-fluorocytosine (5-FC). Cytosine deaminases catalyze the deamination of cytosine to give uracil (Kilstrup M et al. (1989) J Bacteriol 171:2124-2127; Anderson L et al. (1989) Arch Microbiol 152:115-118). Bacteria and fungi which have CDase activity convert 5-FC to the toxic metabolite ("Y") 5-fluorouracil (5-FU) (Polak A & Scholer HJ (1975) Chemotherapy (Basel) 21:113-130). 5-FC itself has low toxicity (Bennett JE, in Goodman and Gilman: the Pharmacological Basis

of Therapeutics. 8th ed., eds. Gilman AG et al. (Pergamon Press, New York) pp. 1165-1181). However, 5-FU has a highly cytotoxic effect, since it is subsequently metabolized to fluoro-UTP (FUTP) and fluoro-dUMP (FdUMP) and thus inhibits RNA and DNA synthesis (Calabrisi P & Chabner BA in Goodman and Gilman: the Pharmacological Basis of Therapeutics. 8th ed., eds. Gilman AG et al. (Pergamon Press, New York) pp. 1209-1263); Damon LE et al. (1989) Pharmac Ther 43:155-189).

Cells of higher plants and mammalian cells have no significant CDase activity and cannot deaminase 5-FC (Polak A et al. (1976) Chemotherapy 22:137-153; Koechlin BA et al. (1966) Biochemical Pharmacology 15:434-446). In this respect, the CDase is introduced as a transgene (e.g. in the form of a transgenic expression cassette) into plant organisms in the course of the process of the invention. Corresponding transgenic plant cells or organisms are then used as masterplants as starting material. Appropriate CDase sequences, transgenic plant organisms and the process of carrying out negative selection processes using, for example, 5-FC as nontoxic substance X, are known to the skilled worker (WO 93/01281; US 5,358,866; Gleave AP et al. (1999) Plant Mol Biol 40(2):223-35; Perera RJ et al. (1993) Plant Mol Biol 23(4):793-799; Stougaard J (1993) Plant J 3:755-761); EP-A1 595 837; Mullen CA et al. (1992) Proc Natl Acad Sci USA 89(1):33-37; Kobayashi T et al. (1995) Jpn J Genet 70(3):409-422; Schlaman HRM & Hooykaas PFF (1997) Plant J 11:1377-1385; Xiaohui Wang H et al. (2001) Gene 272(1-2): 249-255; Koprek T et al. (1999) Plant J 19(6):719-726; Gleave AP et al. (1999) Plant Mol Biol 40(2):223-235; Gallego ME (1999) Plant Mol Biol 39(1):83-93; Salomon S & Puchta H (1998) EMBO J 17(20):6086-6095; Thykjaer T et al. (1997) Plant Mol Biol 35(4):523-530; Serino G (1997) Plant J 12(3):697-701; Risseeuw E (1997) Plant J 11(4):717-728; Blanc V et al. (1996) Biochimie 78(6):511-517; Corneille S et al. (2001) Plant J 27:171-178). Cytosine deaminases and the genes coding therefor may be obtained from a multiplicity of organisms, preferably microorganisms such as, for example, the fungi Cryptococcus neoformans, Candida albicans, Torulopsis glabrata, Sporothrix schenckii, Aspergillus, Cladosporium and Phialophora (JE Bennett, Chapter 50: Antifungal Agents, in Goodman and Gilman's the Pharmacological Basis of Therapeutics 8th ed., A.G. Gilman, ed., Pergamon Press, New York, 1990) and the bacteria E.coli and Salmonella typhimurium (Andersen L et al. (1989) Arch Microbiol 152:115-118).

The sequences, materials and processes disclosed in the context of said publications are hereby explicitly referred to.

Particular preference is given to sequences according to GenBank Acc. No: S56903, and to the modified codA sequences described in EP-A1 595 873, which make expression in eukaryotes possible. Preference is given here to nucleic acid sequences coding for polypeptides according to SEQ ID NO: 2 or, preferably, 4, in particular the sequences according to SEQ ID NO: 1 or, preferably, 3.

(b) cytochrome P-450 enzymes, in particular the bacterial cytochrome P-450 SU1 gene product (CYP105A1) from Streptomyces griseolus (strain ATCC 11796), with preference being given to using as nontoxic substance X substances such as the pro sulfonylurea herbicide R7402 (2-methylethyl-2-3-dihydro-N-[(4,6-dimethoxypyrimidin-2-yl)aminocarbonyl]-1,2-benzoisothiazole-7-sulfonamide 1,1-dioxide). Corresponding sequences and the process of carrying out negative selection processes using, for example, R7402 as nontoxic substance X are known to the skilled worker (O'Keefe DP et al. (1994) Plant Physiol 105:473-482; Tissier AF et al. (1999) Plant Cell 11:1841-1852; Koprek T et al. (1999) Plant J 19(6):719-726; O'Keefe DP (1991) Biochemistry 30(2):447-55). The sequences, materials and processes disclosed in the context of said publications are hereby explicitly referred to.

Particular preference is given to sequences according to GenBank Acc. No: M32238. Preference is further given to nucleic acid sequences coding for the polypeptide according to SEQ ID NO: 6, in particular the sequence according to SEQ ID NO: 5.

(c) indoleacetic acid hydrolases such as, for example, Agrobacterium tumefaciens, tms2 gene product, with preference being given to using as nontoxic substance X substances such as auxin amide compounds or naphthaleneacetamide (NAM) (with NAM being converted to naphthaleneacetic acid, a phytotoxic substance). Corresponding sequences and the process of carrying out negative selection processes using, for example, NAM as nontoxic substance X are known to the skilled worker (Fedoroff NV & Smith DL (1993) Plant J 3:273-289; Upadhyaya NM et al. (2000) Plant Mol Biol Rep 18:227-223; Depicker AG et al. (1988) Plant Cell rep 104:1067-1071; Karlin-Neumannn GA et al. (1991) Plant Cell 3:573-582; Sundaresan V et al. (1995) Gene Develop 9:1797-1810; Cecchini E et al. (1998) Mutat Res 401(1-2):199-206; Zubko E et al. (2000) Nat Biotech-

nol 18:442-445). The sequences, materials and processes disclosed in the context of said publications are hereby explicitly referred to.

Particular preference is given to sequences according to GenBank Acc. No: NC_003308 (Protein_id="NP_536128.1), AE009419, AB016260 (Protein_id="BAA87807.1) and NC002147. Preference is further given to nucleic acid sequences coding for polypeptides according to SEQ ID NO: 8 or 10, in particular the sequences according to SEQ ID NO: 7 or 9.

(d) haloalkane dehalogenases (dhlA gene product), for example from Xanthobacter autotropicus GJ10. The dehalogenase hydrolyzes dihaloalkanes such as 1,2-dichloroethane (DCE) to give halogenated alcohols and inorganic halides (Naested H et al. (1999) Plant J 18(5)571-576; Janssen DB et al. (1994) Annu Rev Microbiol 48: 163-191; Janssen DB (1989) J Bacteriol 171(12):6791-9). The sequences, materials and processes disclosed in the context of said publications are hereby explicitly referred to.

Particular preference is given to sequences according to GenBank Acc. No: M26950. Preference is further given to nucleic acid sequences coding for the polypeptide according to SEQ ID NO: 12, in particular the sequence according to SEQ ID NO: 11.

(e) thymidine kinases (TK), in particular viral TKs from viruses such as Herpes simplex virus, SV40, cytomegalovirus, Varicella zoster virus, in particular the TK of Herpes simplex virus type 1 (TK HSV-1), with preference being given to using as nontoxic substance X substances such as Acyclovir, Ganciclovir or 1,2-deoxy-2-fluoro-β-D-arabinofuranosil-5-iodouracil (FIAU). Corresponding sequences and the process of carrying out negative selection processes using, for example, Acyclovir, Ganciclovir or FIAU as nontoxic substance X are known to the skilled worker (Czako M & Marton L (1994) Plant Physiol 104:1067-1071; Wigler M et al. (1977) Cell 11(1):223-232; McKnight SL et al. (1980) Nucl Acids Res 8(24):5949-5964; McKnight SL et al. (1980) Nucl Acids Res 8(24):5931-5948; Preston et al. (1981) J Virol 38(2):593-605; Wagner et al. (1981) Proc Natl Acad Sci USA 78(3):1441-1445; St. Clair et al.(1987) Antimicrob Agents Chemother 31(6):844-849). The sequences, materials and processes disclosed in the context of said publications are hereby explic-

itly referred to.

Particular preference is given to sequences according to GenBank Acc. No: J02224, V00470 and V00467. Preference is also given to nucleic acid sequences coding for polypeptides according to SEQ ID NO: 14 or 16, in particular the sequences according to SEQ ID NO: 13 or 15.

(f) guanine phosphoribosyl transferases, hypoxanthine phosphoribosyl transferases or xanthine guanine phosphoribosyl transferases, with preference being given to using as nontoxic substance X substances such as 6-thioxanthine or allopurinol. Preference is given to guanine phosphoribosyl transferases (gpt), for example from E. Coli (Besnard et al. (1987) Mol Cell Biol 7:4139; Mzoz and Moolten (1993) Human Gene Therapy 4:589-595; Ono et al. (1997) Hum Gene Ther 8(17):2043-55), hypoxanthine phosphoribosyl transferases (HPRT; Jolly et al. (1983) Proc Natl Acad Sci USA 80:477; Fonwick "The HGPRT Systern", pp. 333-373, M. Gottesman (ed.), Molecular Cell Genetics, John Wiley and Sons, New York, 1985), xanthine guanine phosphoribosyl transferases, for example from Toxoplasma gondii (Knoll LJ et al.(1998) Mol Cell Biol 18(2):807-814; Donald RG et al. (1996) J Biol Chem 271(24):14010-14019). The sequences, materials and processes disclosed in the context of said publications are hereby explicitly referred to.

Particular preference is given to sequences according to GenBank Acc. No: U10247 (Toxoplasma gondii HXGPRT), M13422 (E. coli gpt) and X00221 (E. coli gpt). Preference is also given to nucleic acid sequences coding for polypeptides according to SEQ ID NO: 18, 20 or 22, in particular the sequences according to SEQ ID NO: 17, 19 or 21.

(g) purine nucleoside phosphorylases (PNP; DeoD gene product), for example from E. coli, with preference being given to using as nontoxic substance X substances such as 6-methylpurine deoxyribonucleoside. Corresponding sequences and the process of carrying out negative selection processes using, for example, 6-methylpurine deoxyribonucleoside as nontoxic substance X are known to the skilled worker (Sorscher EJ et al. (1994) Gene Therapy 1:233-238). The sequences, materials and processes disclosed in the context of said publications are hereby explicitly referred to.

Particular preference is given to sequences according to GenBank Acc. No: M60917. Preference is also given to nucleic

acid sequences coding for the polypeptide according to SEQ ID NO: 24, in particular the sequence according to SEQ ID NO: 23.

h) phosphonate monoester hydrolases which convert inactive ester derivatives of the herbicide glyphosate (e.g. glycerylglyphosate) into the active form of the herbicide. Corresponding sequences and the process of carrying out negative selection processes using, for example, glycerylglyphosate are known to the skilled worker (US 5,254,801; Dotson SB et al. (1996) Plant J 10(2):383-392; Dotson SB et al. (1996) J Biol Chem 271(42): 25754-25761). The sequences, materials and processes disclosed in the context of said publications are hereby explicitly referred to.

Particular preference is given to sequences according to GenBank Acc. No: U44852. Preference is also given to nucleic acid sequences coding for the polypeptide according to SEQ ID NO: 26, in particular the sequence according to SEQ ID NO: 25.

(i) aux-1 and, preferably, aux-2 gene products, for example of the Ti plasmids of Agrobacterium strains such as A.rhizogenes or A.tumefaciens (Beclin C et al. (1993) Transgenics Res 2:4855); Gaudin V, Jouanin L. (1995) Plant Mol Biol. 28(1):123-36.

The activity of the two enzymes causes the plant cell to produce indoleacetamide (IAA). Aux-1 encodes an indoleacetamide synthase (IAMS) and converts tryptophan into indoleacetamide (VanOnckelen et al. (1986) FEBS Lett. 198: 357-360). Aux-2 encodes the enzyme indoleacetamide hydrolase (IAMH) and converts indoleacetamide, a substance without phytohormone activity, into the active auxin indoleacetic acid (Inze D et al. (1984) Mol Gen Genet 194:265-274; Tomashow et al. (1984) Proc Natl Acad Sci USA 81:5071-5075; Schroder et al. (1984) Eur J Biochem 138:387-391). The enzyme IAMH may also hydrolyze a number of indoleamide substrates such as, for example, naphthaleneacetamide, the latter being converted into the plant growth regulator naphthaleneacetic acid (NAA). The use of the IAMH gene as a negative selection marker is described, for example, in US 5,180,873. Corresponding enzymes have also been described in A. rhizogenes, A. vitis (Canaday J et al. (1992) Mol Gen Genet 235:292-303) and Pseudomonas savastanoi (Yamada et al. (1985) Proc Natl Acad Sci USA 82:6522-6526). The use as a negative selection marker for destroying partic-

ular cell tissues (e.g. pollen; US 5,426,041) or transgenic plants (US 5,180,873) has been described. Corresponding sequences and the process of carrying out negative selection processes using, for example, naphthaleneacetamide are known to the skilled worker (see above). The sequences, materials and processes disclosed in the context of said publications are hereby explicitly referred to.

Particular preference is given to sequences according to the GenBank Acc. No: M61151, AF039169 and AB025110. Preference is also given to nucleic acid sequences coding for polypeptides according to SEQ ID NO: 28, 30, 32, 34 or 36, in particular the sequences according to SEQ ID NO: 27, 29, 31, 33 or 35.

(j) adenine phosphoribosyl transferases (APRT), with preference being given to using as nontoxic substance X substances such as 4-aminopyrazolopyrimidine. Corresponding sequences and the process of carrying out negative selection processes with use are known to the skilled worker (Wigler M et al. (1979) Proc Natl Acad Sci USA 76(3):1373-6; Taylor et al. "The APRT System", pp., 311-332, M. Gottesman (ed.), Molecular Cell Genetics, John Wiley and Sons, New York, 1985).

k) methoxinine dehydrogenases, with preference being given to using as nontoxic substance X substances such as 2-amino-4-methoxybutanoic acid (methoxinine) which is converted into the toxic methoxyvinyl glycine (Margraff R et al. (1980) Experimentia 36: 846).

l) rhizobitoxin synthases, with preference being given to using as nontoxic substance X substances such as 2-amino-4-methoxybutanoic acid (methoxinine) which is converted into the toxic 2-amino-4-[2-amino-3-hydroxypropyl]-trans-3-butanoic acid (rhizobitoxin) (Owens LD et al. (1973) Weed Science 21:63-66),

m) 5-methylthioribose (MTR) kinases, with preference being given to using as nontoxic substance X substances such as 5-(trifluoromethyl)thioribose (MTR analog, "subversive substrate") which is converted, via an unstable intermediate, into the toxic substance (Y) carbothionyl difluoride. The MTR kinase is a key enzyme of the methionine salvage pathway. Corresponding enzyme activities have been described in plants, bacteria and protozoa but not in mammals. MTR kinases of various species have been identified owing to defined sequence motifs (Sekowska A et al. (2001) BMC Microbiol 1:15;

http://www.biomedcentral.com/1471-2180/1/15). Corresponding sequences and the process of carrying out negative selection processes using, for example, 5-(trifluoromethyl)thioribose are known to the skilled worker and readily obtainable from the appropriate sequence database (e.g. GenBank) (Sekowska A et al. (2001) BMC Microbiol 1:15; Cornell KA et al. (1996) 317:285-290). The sequences, materials and processes disclosed in the context of said publications are hereby explicitly referred to.

However, a plant MTR kinase has not yet been identified unambiguously and is provided within the scope of the process of the invention (SEQ ID NO: 39 and, respectively, 40). In addition, homologs from other plant species are provided, namely from corn (SEQ ID NO: 59 and, respectively, 60), oilseed rape (SEQ ID NO: 61, 63 and, respectively, 62, 64), rice (SEQ ID NO: 65 and, respectively, 66) and soybean (SEQ ID NO: 67 and, respectively, 68).

Accordingly, the invention further relates to amino acid sequences encoding a plant 5-methylthioribose kinase, wherein said amino acid sequence contains at least one sequence selected from the group consisting of SEQ ID NO: 60, 62, 64, 66 or 68.

Accordingly, the invention further relates to nucleic acid sequences encoding a plant 5-methylthioribose kinase, wherein said nucleic acid sequence contains at least one sequence selected from the group consisting of SEQ ID NO: 59, 61, 63, 65 or 67. Even if said sequences are in parts only fragments of complete cDNAs, their length is nevertheless more than sufficient in order to ensure a use and functionality as antisense RNA or double-stranded RNA. Preference is given to using as marker protein a plant endogenous MTR kinase. Further endogenous plant MTR kinases may readily be identified by means of screening databases or gene libraries using conserved, MTK kinase-typical motifs. Said motifs may be derived from Fig. 9a-b, for example. Such motifs may comprise, by way of example but not by limitation, the following sequences:

E(V/I)GDGN(L/I)N(L/Y/F)V(F/Y), preferably EVGDGNLN(Y/F)V(F/Y)
KQALPY(V/I)RC
SWPMT(R/K)ERAYF
PEVYHFDRT
GMRY(I/L)EPPHI
CRLTEQVVFSDPY
HGDLH(S/T)GS

Further suitable motifs may be derived from Fig. 9a-b without difficulty.

Particular preference is given to sequences according to GenBank Acc. No: AF212863 or AC079674 (Protein_ID=AAG51775.1). Preference is also given to nucleic acid sequences coding for polypeptides according to SEQ ID NO: 38 or 40, in particular the sequences according to SEQ ID NO: 37 or 39.

n) alcohol dehydrogenases (Adh), in particular plant Adh-1 gene products, with preference being given to using as nontoxic substance X substances such as allyl alcohol which is converted in this manner into the toxic substance (Y) acrolein. Corresponding sequences and the process of carrying out negative selection processes using, for example, allyl alcohol are known to the skilled worker and readily obtainable from the appropriate sequence database (e.g. GenBank) (Wisman E et al. (1991) Mol Gen Genet 226(1-2):120-8; Jacobs M et al. (1988) Biochem Genet 26(1-2):105-22; Schwartz D. (1981) Environ Health Perspect 37:75-7). The sequences, materials and processes disclosed in the context of said publications are hereby explicitly referred to.

Particular preference is given to sequences according to GenBank Acc. No: X77943, M12196, AF172282, X04049 or AF253472. Preference is also given to nucleic acid sequences coding for polypeptides according to SEQ ID NO: 42, 44, 46 or 48, in particular the sequences according to SEQ ID NO: 41, 43, 45 or 47.

(o) Further suitable negative selection markers are those sequences which exert per se a toxic action on plant cells, such as, for example, diphtheria toxin A, ribonucleases such as barnase and also ribosome-inhibiting proteins such as ricin. In this context, these proteins are preferably expressed in the plant cells inducibly rather than constitutively. The induction is preferably carried out chemically, it being possible, for example, to use the chemically inducible promoters mentioned below in order to ensure said chemically induced expression.

"Reduction" or "to reduce" is to be interpreted broadly in connection with a marker protein or with its amount, expression, activity and/or function and comprises the partial or essentially complete stopping or blocking, based on different cell-biological

mechanisms, of the functionality of a marker protein in a plant cell, plant or a part, tissue, organ, cells or seeds derived therefrom.

A reduction for the purpose of the invention also comprises a reduction of the amount of a marker protein down to an essentially complete lack of said marker protein (i.e. a lack of detectability of marker protein activity or marker protein function or a lack of immunological detectability of said marker protein). In this context, expression of a particular marker protein (or of its amount, expression, activity and/or function) in a cell or an organism is reduced preferably by more than 50%, particularly preferably by more than 80%, very particularly preferably by more than 90%, most preferably by more than 98%. Reduction means in particular also the complete lack of the marker protein (or of its amount, expression, activity and/or function). In this context, activity and/or function mean preferably the property of the marker protein of exerting a toxic effect on the plant cell or the plant organism and, respectively, the ability to convert the substance X into the substance Y. The toxic effect caused by the marker protein is reduced preferably by more than 50%, particularly preferably by more than 80%, very particularly preferably by more than 90%, most preferably by more than 98%. "Reduction" includes of course within the scope of the present invention also a complete, 100% reduction or removal of the marker protein (or of its amount, expression, activity and/or function) (for example by deleting the marker protein gene from the genome).

The invention comprises various strategies for reducing the expression, amount, activity and/or function of the marker protein. The skilled worker appreciates the fact that a number of various methods are available in order to influence the expression, amount, activity and/or function of a marker protein in the desired way. Examples which may be mentioned but which are not limiting are:

a) introducing at least one marker protein double-stranded ribonucleic acid sequence (MP-dsRNA) or an expression cassette or expression cassettes ensuring expression thereof. Included are those processes in which the MP-dsRNA is directed against a marker protein gene (i.e. genomic DNA sequences such as promoter sequences) or a marker protein gene transcript (i.e. mRNA sequences).

b) introducing at least one marker protein antisense ribonucleic acid sequence (MP-antisenseRNA) or an expression cassette ensuring expression thereof. Included are those processes in which the MP-antisenseRNA is directed against a marker protein gene (i.e. genomic DNA sequences) or a marker protein gene transcript (i.e. RNA sequences). α-anomeric nucleic acid sequences are also included.

c) introducing at least one MP-antisenseRNA combined with a ribozyme or an expression cassette ensuring expression thereof

d) introducing at least one marker protein sense ribonucleic acid sequence (MP-senseRNA) for inducing a cosuppression or an expression cassette ensuring expression thereof

e) introducing at least one DNA- or protein-binding factor against a marker protein gene, marker protein RNA or marker protein or an expression cassette ensuring expression thereof

f) introducing at least one viral nucleic acid sequence causing degradation of the marker protein RNA or an expression cassette ensuring expression thereof

g) introducing at least one construct for generating a functional loss (e.g. generation of stop codons, shifts in the reading frame etc.) on a marker protein gene, for example by generating an insertion, deletion, inversion or mutation in a marker protein gene. Preferably, knockout mutants may be generated by means of targeted insertion into said marker protein gene via homologous recombination or by introducing sequence-specific nucleases against marker protein gene sequences.

It is known to the skilled worker that it is also possible to use other processes within the scope of the present invention in order to reduce a marker protein or its activity or function. For example, it may also be advantageous, depending on the type of the marker protein used, to introduce a dominant-negative variant of a marker protein or an expression cassette ensuring expression thereof. In this context, any single one of these processes may cause a reduction in the expression, amount, activity and/or function of a marker protein. A combined application is also conceivable. Further methods are known to the skilled worker and may comprise hindering or stopping the processing of the marker protein, the transport of the marker protein or of its mRNA, the inhibition of ribosome attachment, the inhibition of RNA splicing,

the induction of an enzyme degrading marker protein RNA and/or the inhibition of translational elongation or termination.

The embodiments below will describe by way of example the individual preferred processes:

a) Introducing a double-stranded ribonucleic acid sequence of a marker protein (MP-dsRNA)

The process of gene regulation by means of double-stranded RNA ("double-stranded RNA interference"; dsRNAi) has been described many times for animal and plant organisms (e.g. Matzke MA et al. (2000) Plant Mol Biol 43:401-415; Fire A. et al (1998) Nature 391:806-811; WO 99/32619; WO 99/53050; WO 00/68374; WO 00/44914; WO 00/44895; WO 00/49035; WO 00/63364). The processes and methods described in the references indicated are hereby explicitly referred to. dsRNAi processes are based on the phenomenon that simultaneously introducing the complementary strand and contour strand of a gene transcript suppresses expression of the corresponding gene in a highly efficient manner. Preferably, the phenotype caused is very similar to that of a corresponding knockout mutant (Waterhouse PM et al. (1998) Proc Natl Acad Sci USA 95:13959-64). The dsRNAi process has proved to be particularly efficient and advantageous in reducing marker protein expression.

Double-stranded RNA molecule means within the scope of the invention preferably one or more ribonucleic acid sequences which, owing to complementary sequences, are theoretically (e.g. according to the base pair rules by Watson and Crick) and/or actually (e.g. owing to hybridization experiments in vitro and/or in vivo) capable of forming double-stranded RNA structures. The skilled worker is aware of the fact that the formation of double-stranded RNA structures represents a state of equilibrium. Preferably, the ratio of double-stranded molecules to corresponding dissociated forms is at least 1 to 10, preferably 1:1, particularly preferably 5:1, most preferably 10:1.

The invention therefore further relates to double-stranded RNA molecules (dsRNA-Moleküle) which, when introduced into a plant organism (or into a cell, tissue, organ or propagation material derived therefrom) cause the reduction of at least one marker protein. The double-stranded RNA molecule for reducing expression of a marker protein (MP-dsRNA) here preferably comprises

a) a "sense" RNA strand comprising at least one ribonucleotide sequence which is essentially identical to at least a part of

the "sense" RNA transcript of a nucleic acid sequence coding for a marker protein, and

b) an "antisense" RNA strand which is essentially, preferably fully, complementary to the RNA sense strand under a).

With respect to the dsRNA molecules, marker protein nucleic acid sequence preferably means a sequence according to SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45 or 47 or a functional equivalent thereof.

"Essentially identical" means that the dsRNA sequence may also have insertions, deletions and also individual point mutations in comparison with the marker protein target sequence and nevertheless causes an efficient reduction in expression. The homology (as defined hereinbelow) between the "sense" strand of an inhibitory dsRNA and at least one part of the "sense" RNA transcript of a nucleic acid sequence coding for a market protein (or between the "antisense" strand of the complementary strand of a nucleic acid sequence coding for a marker protein) is preferably at least 75%, preferably at least 80%, very particularly preferably at least 90%, most preferably 100%.

A 100% sequence identity between dsRNA and a marker protein gene transcript is not absolutely necessary in order to cause an efficient reduction in marker protein expression. Consequently, the process is advantageously tolerant toward sequence deviations as may be present due to genetic mutations, polymorphisms or evolutionary divergences. Thus it is possible, for example, using the dsRNA which has been generated starting from the marker protein sequence of the first organism, to suppress marker protein expression in a second organism. This is particularly advantageous when the marker protein used is a plant-intrinsic, endogenous marker protein (for example a 5-methylthioribose kinase or alcohol dehydrogenase). For this purpose, the dsRNA preferably includes sequence regions of marker protein gene transcripts which correspond to conserved regions. Said conserved regions may be readily derived from sequence comparisons.

The length of the subsection is at least 10 bases, preferably at least 25 bases, particularly preferably at least 50 bases, very particularly preferably at least 100 bases, most preferably at least 200 bases or at least 300 bases.

Alternatively, an "essentially identical" dsRNA may also be defined as a nucleic acid sequence capable of hybridizing with part of a marker protein gene transcript (e.g. in 400 mM NaCl, 40 mM

PIPES pH 6.4, 1 mM EDTA at 50°C or 70°C for 12 to 16 h).

"Essentially complementary" means that the "antisense" RNA strand may also have insertions, deletions and also individual point mutations in comparison with the complement of this "sense" RNA strand. The homology between the "antisense" RNA strand and the complement of the "sense" RNA strand is preferably at least 80%, preferably at least 90%, very particularly preferably at least 95%, most preferably 100%.

"Part of the "sense" RNA transcript" of a nucleic acid sequence coding for a marker protein means fragments of an RNA or mRNA transcribed or transcribable from a nucleic acid sequence coding for a marker protein, preferably from a marker protein gene. In this context, the fragments have a sequence length of preferably at least 20 bases, preferably at least 50 bases, particularly preferably at least 100 bases, very particularly preferably at least 200 bases, most preferably at least 500 bases. The complete transcribable RNA or mRNA is also included. Included are also sequences such as those which may be transcribed under artificial conditions from regions of a marker protein gene which are otherwise, under natural conditions, not transcribed, such as promoter regions, for example.

The dsRNA may consist of one or more strands of polyribonucleotides. Naturally, in order to achieve the same purpose, it is also possible to introduce a plurality of individual dsRNA molecules which comprise in each case one of the above-defined ribonucleotide sequence sections into the cell or the organism. The double-stranded dsRNA structure may be formed starting from two complementary, separate RNA strands or, preferably, starting from a single, self-complementary RNA strand. In this case, the "sense" RNA strand and the "antisense" RNA strand are preferably connected covalently to one another in the form of an inverted "repeat".

As described in WO 99/53050, for example, the dsRNA may also comprise a hairpin structure by connecting the "sense" and the "antisense" strands by a connecting sequence ("linker"; for example an intron). Preference is given to the self-complementary dsRNA structures, since they require only the expression of an RNA sequence and always comprise the complementary RNA strands in an equimolar ratio. The connecting sequence may is preferably an intron (e.g. an intron of the potato ST-LS1 gene; Vancanneyt GF et al. (1990) Mol Gen Genet 220(2):245-250).

The nucleic acid sequence coding for a dsRNA may include further elements such as, for example, transcription termination signals or polyadenylation signals.

Bringing together, if intended, the two strands of the dsRNA in a cell or plant may be achieved by way of example in the following way:

a) transformation of the cell or plant with a vector comprising both expression cassettes,

b) cotransformation of the cell or plant with two vectors, one of which comprises the expression cassettes containing the "sense" strand and the other one of which comprises the expression cassettes containing the "antisense" strand.

The formation of the RNA duplex may be initiated either outside or inside the cell.

The dsRNA may be synthesized either in vivo or in vitro. For this purpose, a DNA sequence coding for a dsRNA may be inserted into an expression cassette under the control of at least one genetic control element (such as a promoter, for example). A polyadenylation is not necessary and neither need any elements for initiating a translation be present. Preference is given to the expression cassette for the MP-dsRNA being present on the transformation construct or the transformation vector. For this purpose, the expression cassettes coding for the "antisense" strand and/or the "sense" strand of an MP-dsRNA or for the self-complementary strand of the dsRNA are preferably inserted into a transformation vector and introduced into the plant cell by using the processes described below. A stable insertion into the genome may be advantageous for the process of the invention but is not absolutely necessary. Since a dsRNA causes a long-term effect, transient expression is also sufficient in many cases. The dsRNA may also be part of the RNA to be expressed by the nucleic acid sequence to be inserted by fusing it, for example, to the 3'-untranslated part of said RNA.

The dsRNA may be introduced in an amount which makes possible at least one copy per cell. Higher amounts (e.g. at least 5, 10, 100, 500 or 1000 copies per cell) may, if appropriate, cause a more efficient reduction.

b) Introducing an antisense ribonucleic acid sequence of a marker protein (MP-antisenseRNA)

Processes for reducing a particular protein by means of the "antisense" technique have been described multiple times, also in plants (Sheehy et al. (1988) Proc Natl Acad Sci USA 85: 8805-8809; US 4,801,340; Mol JN et al. (1990) FEBS Lett 268(2):427-430). The antisense nucleic acid molecule hybridizes or binds to the cellular mRNA and/or genomic DNA coding for the marker protein to be reduced, thereby suppressing transcription and/or translation of said marker protein. The hybridization may be produced in a conventional manner via the formation of a stable duplex or, in the case of genomic DNA, by binding of the antisense nucleic acid molecule to the duplex of the genomic DNA via specific interaction in the large groove of the DNA helix.

An MP-antisenseRNA may be derived using the nucleic acid sequence coding for this marker protein, for example the nucleic acid sequence according to SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45 or 47 according to the base pair rules by Watson and Crick. The MP-antisenseRNA may be complementary to the entire transcribed mRNA of the marker protein, may be limited to the coding region or may consist only of an oligonucleotide which is complementary to a part of the coding or noncoding sequence of the mRNA. Thus, for example, the oligonucleotide may be complementary to the region comprising the translation start site for the marker protein. The MP-antisenseRNA may be, for example, 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length, but may also be longer and comprise at least 100, 200, 500, 1000, 2000 or 5000 nucleotides. MP-antisenseRNA are preferably expressed recombinantly in the target cell in the course of the process of the invention.

The MP-antisenseRNA may also be part of an RNA to be expressed by the nucleic acid sequence to be inserted by being fused, for example, to the 3'-untranslated part of said RNA.

The invention further relates to transgenic expression cassettes containing a nucleic acid sequence coding for at least part of a marker protein, with said nucleic acid sequence being functionally linked in antisense orientation to a promoter functional in plant organisms. Said expression cassettes may be part of a transformation construct or transformation vector or else may be introduced in the course of a cotransformation.

In a further preferred embodiment, expression of a marker protein may be inhibited by nucleotide sequences which are complementary to the regulatory region of a marker protein gene (e.g. a marker protein promoter and/or enhancer) and which form with the DNA double helix there triple-helical structures, thereby reducing transcription of the marker protein gene. Corresponding processes

have been described (Helene C (1991) Anticancer Drug Res 6(6):569-84; Helene C et al. (1992) Ann NY Acad Sci 660:27-36; Maher LJ (1992) Bioassays 14(12):807-815).

In a further embodiment, the MP-antisenseRNA may be an α-anomeric nucleic acid. Such α-anomeric nucleic acid molecules form with complementary RNA specific double-stranded hybrids in which, in contrast to the conventional β-nucleic acids, the two strands are oriented parallel to one another (Gautier C et al. (1987) Nucleic Acids Res 15:6625-6641).

c) Introducing an MP-antisenseRNA combined with a ribozyme

Advantageously, the above-described antisense strategy may be coupled to a ribozyme process. Catalytic RNA molecules or ribozymes may be adapted to any target RNA and cleave the phosphodiester backbone in specific positions, thereby functionally deactivating said target RNA (Tanner NK (1999) FEMS Microbiol Rev 23(3):257-275). In the process, the ribozyme is not modified itself but is capable of cleaving in an analogous manner further target RNA molecules, thereby acquiring the properties of an enzyme. The incorporation of ribozyme sequences into "antisense" RNAs imparts specifically to these "antisense" RNAs this enzyme-like, RNA-cleaving property and thus increases their efficiency in inactivating the target RNA. The preparation and use of appropriate ribozyme "antisense" RNA molecules have been described (inter alia in Haseloff et al. (1988) Nature 334: 585-591); Haselhoff and Gerlach (1988) Nature 334:585-591; Steinecke P et al. (1992) EMBO J 11(4):1525- 1530; de Feyter R et al. (1996) Mol Gen Genet. 250(3):329-338).

In this way, it is possible to use ribozymes (e.g. hammerhead ribozymes; Haselhoff and Gerlach (1988) Nature 334:585-591) in order to catalytically cleave the mRNA of a marker protein to be reduced and thus prevent translation. The ribozyme technique may increase the efficiency of an antisense strategy. Processes for expressing ribozymes in order to reduce particular proteins have been described in (EP 0 291 533, EP 0 321 201, EP 0 360 257). Ribozyme expression has likewise been described in plant cells (Steinecke P et al. (1992) EMBO J 11(4):1525-1530; de Feyter R et al. (1996) Mol Gen Genet. 250(3):329-338). Suitable target sequences and ribozymes may be determined, for example, as described in "Steinecke P, Ribozymes, Methods in Cell Biology 50, Galbraith et al. eds, Academic Press, Inc. (1995), pp. 449-460", by calculating the secondary structures of ribozyme RNA and target RNA and by the interaction thereof (Bayley CC et al. (1992) Plant Mol Biol. 18(2):353-361; Lloyd AM and Davis RW et al. (1994) Mol Gen Genet. 242(6):653-657). It is possible, for exam-

ple, to construct derivatives of the Tetrahymena L-19 IVS RNA which have regions complementary to the mRNA of the marker protein to be suppressed (see also US 4,987,071 and US 5,116,742). Alternatively, such ribozymes may also be identified via a selection process from a library of various ribozymes (Bartel D and Szostak JW (1993) Science 261:1411-1418).

d) Introducing a sense ribonucleic acid sequence of a marker protein (MP-senseRNA) for inducing a cosuppression

Expression of a marker protein ribonucleic acid sequence (or a part thereof) in sense orientation may result in a cosuppression of the corresponding marker protein gene. Expression of sense RNA with homology to an endogenous marker protein gene may reduce or switch off expression of the latter, as has been described similarly for antisense approaches (Jorgensen et al. (1996) Plant Mol Biol 31(5):957-973; Goring et al. (1991) Proc Natl Acad Sci USA 88:1770-1774; Smith et al. (1990) Mol Gen Genet 224:447-481; Napoli et al. (1990) Plant Cell 2:279-289; Van der Krol et al. (1990) Plant Cell 2:291-99). In this context, the introduced construct may represent completely or only partially the homologous gene to be reduced. The possibility of translation is not required. The application of this technique to plants has been described (e.g. Napoli et al. (1990) Plant Cell 2:279-289; in US 5,034,323.

The cosuppression is preferably carried out using a sequence which is essentially identical to at least part of the nucleic acid sequence coding for a marker protein, for example the nucleic acid sequence according to SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29 ,31, 33, 35, 37, 39, 41, 43, 45 or 47.

The MP-senseRNA is preferably chosen in such a way that a translation of the marker protein or a part thereof cannot occur. For this purpose, for example, the 5'-untranslated or 3'-untranslated region may be chosen or else the ATG start codon may be deleted or mutated.

e) Introducing DNA- or protein-binding factors against marker protein genes, marker protein RNAs or proteins

Marker protein expression may also be reduced using specific DNA-binding factors, for example factors of the zinc finger transcription factor type. These factors attach to the genomic sequence of the endogenous target gene, preferably in the

regulatory regions, and cause a reduction in expression. Appropriate processes for preparing corresponding factors have been described (Dreier B et al. (2001) J Biol Chem 276(31):29466-78; Dreier B et al. (2000) J Mol Biol 303(4):489-502; Beerli RR et al. (2000) Proc Natl Acad Sci USA 97 (4):1495-1500; Beerli RR et al. (2000) J Biol Chem 275(42):32617-32627; Segal DJ and Barbas CF 3rd. (2000) Curr Opin Chem Biol 4(1):34-39; Kang JS and Kim JS (2000) J Biol Chem 275(12):8742-8748; Beerli RR et al. (1998) Proc Natl Acad Sci USA 95(25):14628- 14633; Kim JS et al. (1997) Proc Natl Acad Sci USA 94(8):3616 -3620; Klug A (1999) J Mol Biol 293(2):215-218; Tsai SY et al. (1998) Adv Drug Deliv Rev 30(1-3):23-31; Mapp AK et al. (2000) Proc Natl Acad Sci USA 97(8):3930-3935; Sharrocks AD et al. (1997) Int J Biochem Cell Biol 29(12):1371-1387; Zhang L et al. (2000) J Biol Chem 275(43):33850-33860).

These factors may be selected using any segment of a marker protein gene. This section is preferably in the region of the promoter region. However, for gene suppression, it may also be in the region of the coding exons or introns.

It is also possible to introduce factors which inhibit the marker protein itself into a cell. These protein-binding factors may be, for example, aptamers (Famulok M and Mayer G (1999) Curr Top Microbiol Immunol 243:123-36) or antibodies or antibody fragments or single-chain antibodies. Obtaining these factors has been described (Owen M et al. (1992) Biotechnology (N Y) 10(7):790-794; Franken E et al. (1997) Curr Opin Biotechnol 8(4):411-416; Whitelam (1996) Trend Plant Sci 1:286-272).

f) Introducing viral nucleic acid sequences and expression constructs causing the degradation of marker protein RNA

Marker protein expression may also be effectively implemented by inducing the specific degradation of marker protein RNA by the plant with the aid of a viral expression system (Amplikon; Angell SM et al. (1999) Plant J 20(3):357-362). These systems, also referred to as "VIGS" (viral induced gene silencing), introduce nucleic acid sequences with homology to the transcript of a marker protein to be reduced into the plant by means of viral vectors. Transcription is then switched off, presumably mediated by plant defence mechanisms against viruses. Appropriate techniques and processes have been described (Ratcliff F et al. (2001) Plant J 25(2):237-45; Fagard M und Vaucheret H (2000) Plant Mol Biol 43(2-3):285-93; Anandalakshmi R et al. (1998) Proc Natl Acad Sci USA 95(22):13079-84; Ruiz MT (1998) Plant Cell 10(6):937-46).

VIGS-mediated reduction is preferably implemented using a sequence which is essentially identical to at least part of the nucleic acid sequence coding for a marker protein, for example the nucleic acid sequence according to SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45 or 47.

g) Introducing constructs for generating a functional loss or a functional reduction of marker protein genes

The skilled worker knows numerous possible processes of how to modify genomic sequences in a targeted manner. These include, in particular, processes such as the generation of knockout mutants by means of targeted homologous recombination, for example by generating stop codons, shifts in the reading frame etc. (Hohn B and Puchta H (1999) Proc Natl Acad Sci USA 96:8321-8323) or the targeted deletion or inversion of sequences by means of, for example, sequence-specific recombinases or nucleases (see below).

In a preferred embodiment, the marker protein gene is inactivated by introducing a sequence-specific recombinase. Thus it is possible, for example, for the marker protein gene to include recognition sequences for sequence-specific recombinases or to be flanked by such sequences, and introducing the recombinase then deletes or inverts particular sequences of the marker protein gene, thus leading to inactivation of the marker protein gene. A corresponding procedure is depicted diagrammatically in Fig. 1.

Appropriate processes for deletion/inversion of sequences by means of sequence-specific recombinase systems are known to the skilled worker. Examples which may be mentioned are the Cre/lox system of bacteriophage P1 (Dale EC and Ow DW (1991) Proc Natl Acad Sci USA 88:10558-10562; Russell SH et al. (1992) Mol Gen Genet 234:49-59; Osborne BI et al. (1995) Plant J 7:687-701), the yeast FLP/FRT system (Kilby NJ et al. (1995) Plant J 8:637-652; Lyznik LA et al. (1996) Nucl Acids Res 24:3784-3789), the Gin recombinase of the Mu phage, the E. coli Pin recombinase and the R/RS system of the pSR1 plasmids (Onouchi H et al.(1995) Mol Gen Genet 247:653-660; Sugita Ket al. (2000) Plant J. 22:461-469). In these systems, the recombinase (for example Cre or FLP) interacts specifically with its particular recombination sequences (34 bp lox-Sequenz and, respectively, 47 bp FRT sequence). Preference is given to the bacteriophage P1 Cre/lox and the yeast FLP/FRT systems. The FLP/FRT and cre/lox recombinase systems have already been applied in plant systems (Odell et al. (1990) Mol Gen Genet 223:369-378). Preference is given to introducing the recombinase

by means of recombinant expression starting from an expression cassette included on a DNA construct.

The activity or amount of the marker protein may also be reduced by a targeted deletion in the marker protein gene, for example by sequence-specific induction of DNA double-strand breaks at a recognition sequence for specific induction of DNA double-strand breaks in or close to the nucleic acid sequence coding for a marker protein. In its simplest embodiment (cf. Fig. 2, A and B) an enzyme is to this end introduced with the transformation construct, which generates at least one double-strand break in such a way that the resulting illegitimate recombination or deletion causes a reduction in the activity or amount of marker protein, for example by inducing a shift in the reading frame or deletion of essential sequences.

The efficiency of this approach may be increased by the sequence coding for the marker protein being flanked by sequences (A and, respectively, A') which have a sufficient length and homology to one another in order to recombine with one another as a consequence of the induced double-strand break and thus to cause, due to an intramolecular homologous recombination, a deletion of the sequence coding for the marker protein. Fig. 3 depicts diagrammatically a corresponding procedure in an exemplary embodiment of this variant.

The amount, function and/or activity of the marker protein may also be reduced by a targeted insertion of nucleic acid sequences (for example of the nucleic acid sequence to be inserted within the scope of the process of the invention) into the sequence coding for a marker protein (e.g. by means of intermolecular homologous recombination). This embodiment of the process of the invention is particularly advantageous and preferred, since, in addition to the general advantages of the process of the invention, it makes it moreover also possible to insert the nucleic acid sequence to be inserted into the plant genome in a reproducible, predictable, location-specific manner. This avoids the positional effects which otherwise occur in the course of a random, location-unspecific insertion (and which may manifest themselves, for example, in the form of different levels of expression of the transgene or in unintended inactivation of endogenous genes). Preference is given to using as an "anti-marker protein" compound in the course of this embodiment a DNA construct which comprises at least part of the sequence of a marker protein gene or neighbouring sequences and which can thus specifically recombine with said sequences in the target cell so that a deletion,

addition or substitution of at least one nucleotide alters the marker protein gene in such a way that the functionality of said marker protein gene is reduced or completely removed. The alteration may also affect the regulatory elents (e.g. the promoter) of the marker protein gene so that the coding sequence remains unaltered, but expression (transcription and/or translation) does not occur and is reduced. In conventional homologous recombination, the sequence to be inserted is flanked at its 5' and/or 3' end by further nucleic acid sequences (A' and, respectively, B') which have a sufficient length and homology to corresponding sequences of the marker protein gene (A and, respectively, B) for making homologous recombination possible. The length is usually in a range from several hundred bases to several kilobases (Thomas KR and Capecchi MR (1987) Cell 51:503; Strepp et al. (1998) Proc Natl Acad Sci USA 95(8):4368-4373). The homologous recombination is carried out by transforming the plant cell containing the recombination construct by using the process described below and selecting successfully recombined clones based on the subsequently inactivated marker protein. Although homologous recombination is a relatively rare event in plant organisms, a selection pressure may be avoided by recombination into the marker protein gene, allowing a selection of the recombined cells and sufficient efficiency of the process. Fig. 4 diagrammatically depicts a corresponding procedure in an exemplary embodiment of this variant.

In an advantageous embodiment of the invention, however, insertion into the marker protein gene is facilitated by means of further functional elements. The term is to be understood as being comprehensive and means the use of sequences or of transcripts or polypeptides derived therefrom which are capable of increasing the efficiency of the specific integration into a marker protein gene. Various processes are available to the skilled worker for this purpose. However, preference is given to implementing the insertion by inducing a sequence-specific double-strand break in or close to the marker protein gene.

In a preferred embodiment of the invention, the marker protein is inactivated (i.e. the amount, expression, activity or function is reduced) by integrating a DNA sequence into a marker protein gene, with the process preferably comprising the following steps:

i) introducing an insertion construct and at least one enzyme suitable for inducing DNA double-strand breaks at a recognition sequence for targeted induction of DNA double-strand

breaks in or close to the marker protein gene, and

ii) inducing DNA double-strand breaks at the recognition sequences for targeted induction of DNA double-strand breaks in or close to the marker protein gene, and

iii) inserting the insertion construct into the marker protein gene, with the functionality of the marker protein gene and, preferably, the functionality of the recognition sequence for targeted induction of DNA double-strand breaks is inactivated so that the enzyme suitable for induction of DNA double-strand breaks can no longer cut said recognition sequence, and

iv) selecting plants or plant cells in which the insertion construct has been inserted into the marker protein gene.

The insertion construct, preferably, comprises the nucleic acid sequence to be inserted into the genome but may also be used separately therefrom.

"Enzyme suitable for inducing DNA double-strand breaks at the recognition sequence for targeted induction of DNA double-strand breaks" ("DSBI enzyme" for "double-strand-break inducing enzyme" hereinbelow) means generally all those enzymes which are capable of generating sequence-specifically double-strand breaks in double-stranded DNA. Examples which may be mentioned but which are not limiting are:

1. Restriction endonucleases, preferably type II restriction endonucleases, particularly preferably Homing endonucleases as described in detail hereinbelow.

2. Artificial nucleases as described in detail hereinbelow, such as, for example, chimeric nucleases, mutated restriction or Homing endonucleases or RNA protein particles derived from group II mobile introns.

Both natural and artificially prepared DSBI enzymes are suitable. Preference is given to all of those DSBI enzymes whose recognition sequence is known and which can either be obtianed in the form of their proteins (for example by purification) or be expressed using their nucleic acid sequence.

Preference is given to selecting the DSBI enzyme, with the knowledge of its specific recognition sequence, in such a way that it

possesses, apart from the target recognition sequence, no further functional recognition regions in the genome of the target plant. Very particular preference is therefore given to Homing endonucleases (overview: Belfort M and Roberts RJ (1997) Nucleic Acids Res 25:3379-3388; Jasin M (1996) Trends Genet 12:224-228; Internet: http://rebase.neb.com/rebase/rebase.homing.html; Roberts RJ and Macelis D (2001) Nucl Acids Res 29: 268-269). The latter fulfill said requirement, owing to their long recognition sequences. The sequences coding for Homing endonucleases of this kind may be isolated, for example, from the Chlamydomonas chromoplast genome (Turmel M et al. (1993) J Mol Biol 232:446-467). Suitable Homing endonucleases are listed under the abovementioned internet address. Examples of Homing endonucleases which may be mentioned are those like F-SceI, F-SceII, F-SuvI, F-TevI, F-TevII, I-AmaI, I-AniI, I-CeuI, I-CeuAIIP, I-ChuI, I-CmoeI, I-CpaI, I-CpaII, I-CreI, I-CrepsbIP, I-CrepsbIIP, I-CrepsbIIIP, I-CrepsbIVP, I-CsmI, I-CvuI, I-CvuAIP, I-DdiII, I-DirI, I-DmoI, I-HspNIP, I-LlaI, I-MsoI, I-NaaI, I-NanI, I-NclIP, I-NgrIP, I-NitI, I-NjaI, I-Nsp236IP, I-PakI, I-PboIP, I-PcuIP, I-PcuAI, I-PcuVI, I-PgrIP, I-PobIP, I-PorI, I-PorIIP, I-PpbIP, I-PpoI, I-SPBetaIP, I-ScaI, I-SceI, I-SceII, I-SceIII , I-SceIV, I-SceV, I-SceVI, I-SceVII, I-SexIP, I-SneIP, I-SpomCP, I-SpomIP, I-SpomIIP, I-SquIP, I-Ssp6803I, I-SthPhiJP, I-SthPhiST3P, I-SthPhiS3bP, I-TdeIP, I-TevI, I-TevII, I-TevIII, I-UarAP, I-UarHGPA1P, I-UarHGPA13P, I-VinIP, I-ZbiIP, PI-MtuI, PI-MtuHIP, PI-MtuHIIP, PI-PfuI, PI-PfuII, PI-PkoI, PI-PkoII, PI-PspI, PI-Rma43812IP, PI-SPBetaIP, PI-SceI, PI-TfuI, PI-TfuII, PI-ThyI, PI-TliI, PI-TliII. Preference is given here to those Homing endonucleases whose gene sequences are already known, such as, for example, F-SceI, I-CeuI, I-ChuI, I-DmoI, I-CpaI, I-CpaII, I-CreI, I-CsmI, F-TevI, F-TevII, I-TevI, I-TevII, I-AniI, I-CvuI, I-LlaI, I-NanI, I-MsoI, I-NitI, I-NjaI, I-PakI, I-PorI, I-PpoI, I-ScaI, I-Ssp6803I, PI-PkoI, PI-PkoII, PI-PspI, PI-TfuI, PI-TliI.

Very particular preference is given to

- I-CeuI (Cote MJ and Turmel M (1995) Curr Genet 27:177-183.; Gauthier A et al. (1991) Curr Genet 19:43-47; Marshall (1991) Gene 104:241-245; GenBank Acc. No.: Z17234 nucleotides 5102 to 5758),

- I-ChuI (Cote V et al.(1993) Gene 129:69-76; GenBank Acc. No.: L06107, nucleotides 419 to 1075),

- I-CmoeI (Drouin M et al. (2000) Nucl Acids Res 28:4566-4572),

- I-CpaI from *Chlamydomonas pallidostigmatica* (GenBank Acc. No.: L36830, nucleotides 357 to 815; Turmel M et al. (1995) Nucleic Acids Res 23:2519-2525; Turmel, M et al. (1995) Mol Biol Evol 12:533-545)

- I-CpaII (Turmel M et al. (1995) Mol Biol Evol 12:533-545; GenBank Acc. No.: L39865, nucleotides 719 to 1423),

- I-CreI (Wang J et al. (1997) Nucleic Acids Res 25: 3767-3776; Dürrenberger, F and Rochaix JD (1991) EMBO J 10:3495-3501; GenBank Acc. No.: X01977, nucleotides 571 to 1062),

- I-CsmI (Ma DP et al. (1992) Plant Mol Biol 18:1001-1004)

- I-NanI (Elde M et al. (1999) Eur J Biochem. 259:281-288; GenBank Acc. No.: X78280, nucleotides 418 to 1155),

- I-NitI (GenBank Acc. No.: X78277, nucleotides 426 to 1163),

- I-NjaI (GenBank Acc. No.: X78279, nucleotides 416 to 1153),

- I-PpoI (Muscarella DE and Vogt VM (1989) Cell 56:443-454; Lin J and Vogt VM (1998) Mol Cell Biol 18:5809-5817; GenBank Acc. No.: M38131, nucleotides 86 to 577),

- I-PspI (GenBank Acc. No.: U00707, nucleotides 1839 to 3449),

- I-ScaI (Monteilhet C et al. (2000) Nucleic Acids Res 28: 1245-1251; GenBank Acc. No.: X95974, nucleotides 55 to 465)

- I-SceI (WO 96/14408; US 5,962,327, therein Seq ID NO: 1),

- Endo SceI (Kawasaki et al. (1991) J Biol Chem 266:5342-5347, identical to F-SceI; GenBank Acc. No.: M63839, nucleotides 159 to 1589),

- I-SceII (Sarguiel B et al. (1990) Nucleic Acids Res 18:5659-5665),

- I-SceIII (Sarguiel B et al. (1991) Mol Gen Genet. 255:340-341),

- I-Ssp6803I (GenBank Acc. No.: D64003, nucleotides 35372 to 35824),

- I-TevI (Chu et al. (1990) Proc Natl Acad Sci USA 87:3574-3578; Bell-Pedersen et al. (1990) Nucleic Acids Res18:3763-3770; GenBank Acc. No.: AF158101, nucleotides 144431 to 143694),

- I-TevII (Bell-Pedersen et al. (1990) Nucleic Acids Res 18:3763-3770; GenBank Acc. No.: AF158101, nucleotides 45612 to 44836),

- I-TevIII (Eddy et al. (1991) Genes Dev. 5:1032-1041).

Very particular preference is given to commercially available Homing endonucleases such as I-CeuI, I-SceI, I-PpoI, PI-PspI or PI-SceI. Most preference is given to I-SceI and I-PpoI. While the gene coding for I-PpoI may be utilized in its natural form, the gene coding for I-SceI possesses an editing site. Since, in contrast to yeast mitochondria, the appropriate editing is not carried out in higher plants, an artificial sequence encoding the I-SceI protein must be used for heterologous expression of this enzyme (US 5,866,361).

The enzymes may be purified from their source organisms in the manner familiar to the skilled worker and/or the nucleic acid sequence encoding said enzymes may be cloned. The sequences of various enzymes have been deposited with GenBank (see above).

Artificial DSBI enzymes which may be mentioned by way of example are chimeric nucleases which are composed of an unspecific nuclease domain and a sequence-specific DNA-binding domain (e.g. consisting of zinc fingers) (Smith J et al. (2000) Nucl Acids Res 28(17):3361-3369; Bibikova M et al. (2001) Mol Cell Biol. 21:289-297). Thus, for example, the catalytic domain of the restriction endonuclease FokI has been fused to zinc finger-binding domains, thereby defining the specificity of the endonuclease (Chandrasegaran S & Smith J (1999) Biol Chem 380:841-848; Kim YG & Chandrasegaran S (1994) Proc Natl Acad Sci USA 91:883-887; Kim YG et al. (1996) Proc Natl Acad Sci USA 93:1156-1160). The described technique has also been used previously for imparting a predefined specificity to the catalytic domain of the yeast Ho endonuclease by fusing said domain to the zinc finger domain of transcription factors (Nahon E & Raveh D (1998) Nucl Acids Res 26:1233-1239). It is possible, using suitable mutation and selection processes, to adapt existing Homing endonucleases to any de-

sired recognition sequence.

As mentioned, zinc finger proteins are particularly suitable as DNA-binding domains within chimeric nucleases. These DNA-binding zinc finger domains may be adapted to any DNA sequence. Appropriate processes for preparing corresponding zinc finger domains have been described and are known to the skilled worker (Beerli RR et al. (2000) Proc Natl Acad Sci 97(4):1495-1500; Beerli RR et al.(2000) J Biol Chem 275(42):32617-32627; Segal DJ and Barbas CF 3rd. (2000) Curr Opin Chem Biol 4(1):34-39; Kang JS and Kim JS (2000) J Biol Chem 275(12):8742-8748; Beerli RR et al. (1998) Proc Natl Acad Sci USA 95(25):14628-14633; Kim JS et al. (1997) Proc Natl Acad Sci USA 94(8):3616-3620; Klug A (1999) J Mol Biol 293(2):215-218; Tsai SY et al. (1998) Adv Drug Deliv Rev 30(1-3):23-31; Mapp AK et al. (2000) Proc Natl Acad Sci USA 97(8):3930-3935; Sharrocks AD et al. (1997) Int J Biochem Cell Biol 29(12):1371-1387; Zhang L et al. (2000) J Biol Chem 275(43):33850-33860). Processes for preparing and selecting zinc finger DNA-binding domains with high sequence specificity have been described (WO 96/06166, WO 98/53059, WO 98/53057). Fusing a DNA-binding domain obtained in this way to the catalytic domain of an endonuclease (such as, for example, the FokI or Ho endonuclease) enables chimeric nucleases to be prepared which have any desired specificity and which may be used as DSBI enzymes advantageously within the scope of the present invention.

Artificial DSBI enzymes with altered sequence specificity may also be generated by mutating already known restriction endonucleases or Homing endonucleases, using methods familiar to the skilled worker. Besides the mutagenesis of Homing endonucleases, the mutagenesis of maturases is of particular interest for the purpose of obtaining an altered substrate specificity. Maturases frequently share many features with Homing endonucleases and, if appropriate, can be converted into nucleases by carrying out few mutations. This has been shown, for example, for the maturase in the bakers' yeast bi2 intron. Only two mutations in the maturase-encoding open reading frame (ORF) sufficed to impart to this enzyme a Homing-endonuclease activity (Szczepanek & Lazowska (1996) EMBO J 15:3758-3767).

Further artificial nucleases may be generated with the aid of mobile group II introns and the proteins encoded by them, or parts of these proteins. Mobile group II introns, together with the proteins encoded by them, form RNA-protein particles which are capable of recognizing and cutting DNA in a sequence-specific manner. In this context, the sequence specificity can be adapted

to the requirements by mutating particular regions of the intron (see below) (WO 97/10362).

Preference is given to expressing the DSBI enzyme as a fusion protein with a nuclear localization sequence (NLS). This NLS sequence enables facilitated transport into the nucleus and increases the efficiency of the recombination system. Various NLS sequences are known to the skilled worker and described, inter alia, in Jicks GR and Raikhel NV (1995) Annu. Rev. Cell Biol. 11:155-188. For example, the NLS sequence of the SV40 large antigen is preferred for plant organisms. Very particular preference is given to the following NLS sequences:

NLS1: N-Pro-Lys-Thr-Lys-Arg-Lys-Val-C

NLS2: N-Pro-Lys-Lys-Lys-Arg-Lys-Val-C

Owing to the small size of many DSBI enzymes (such as, for example, the Homing endonucleases), an NLS sequence is not absolutely necessary, however. These enzymes are able to pass through the nuclear pores also without this assistance.

"Recognition sequence for targeted induction of DNA double-strand breaks" means in general those sequences which allow recognition and cleavage by the DSBI enzyme under the conditions in the eukaryotic cell or organism used in this case. In this context, mention is made, by way of example but not by limitation, in table 1 below of the recognition sequences for the particular DSBI enzymes listed.

Table 1: Recognition sequences and source organisms of DSBI enzymes ("^" indicates the cleavage site of the DSBI enzyme within a recognition sequence)

DSBI enzyme	Source organism	Recognition sequence
CRE	Bacteriophage P1	5'-AACTCTCATCGCTTCGGATAACTTCCTGTTATCCGAAACAT ATCACTCACTTTGGTGATTTCACCGTAACTGTCTATGATTAATG -3'
FLP	Saccharomyces cerevisiae	5'-GAAGTTCCTATTCCGAAGTTCCTATTCTCTAGAAAGTA-TAGGAACTTC-3'
R	pSR1 plasmids	5'-CGAGATCATATCACTGTGGACGTTGATGAAAGAATACGTTA TTCTTTCATCAAATCGT

P-element transposase	Drosophila	5'-CTAGATGAAATAACATAAGGTGG
I-AniI	Aspergillus nidulans	5'-TTGAGGAGGTT^TCTCTGTAAATAANNNNNNNNNNNNNNN 3'-AACTCCTCCAAAGAGACATTTATTNNNNNNNNNNNNNNN^
I-DdiI	Dictyostelium discoideumAX3	5'-TTTTTTGGTCATCCAGAAGTATAT 3'-AAAAAACCAG^TAGGTCTTCATATA
I-CvuI	Chlorella vulgaris	5'-CTGGGTTCAAAACGTCGTGA^GACAGTTTGG 3'-GACCCAAGTTTTGCAG^CACTCTGTCAAACC
I-CsmI	Chlamydomonas smithii	5'-GTACTAGCATGGGGTCAAATGTCTTTCTGG
I-CmoeI	Chlamydomonas moewusii	5'-TCGTAGCAGCT^CACGGTT 3'-AGCATCG^TCGAGTGCCAA
I-CreI	Chlamydomonas reinhardtii	5'-CTGGGTTCAAAACGTCGTGA^GACAGTTTGG 3'-GACCCAAGTTTTGCAG^CACTCTGTCAAACC
I-ChuI	Chlamydomonas humicola	5'-GAAGGTTTGGCACCTCG^ATGTCGGCTCATC 3'-CTTCCAAACCGTG^GAGCTACAGCCGAGTAG
I-CpaI	Chlamydomonas pallidostigmatica	5'-CGATCCTAAGGTAGCGAA^ATTCA 3'-GCTAGGATTCCATC^GCTTTAAGT
I-CpaII	Chlamydomonas pallidostigmatica	5'-CCCGGCTAACTC^TGTGCCAG 3'-GGGCCGAT^TGAGACACGGTC
I-CeuI	Chlamydomonas eugametos	5'-CGTAACTATAACGGTCCTAA^GGTAGCGAA 3'-GCATTGATATTGCCAG^GATTCCATCGCTT
I-DmoI	Desulfurococcus mobilis	5'-ATGCCTTGCCGGGTAA^GTTCCGGCGCGCAT 3'-TACGGAACGGCC^CATTCAAGGCCGCGCGTA
I-SceI	S.cerevisiae	5'-AGTTACGCTAGGGATAA^CAGGGTAATATAG 3'-TCAATGCGATCCC^TATTGTCCCATTATATC 5'-TAGGGATAA^CAGGGTAAT 3'-ATCCC^TATTGTCCCATTA ("Core" sequence)
I-SceII	S.cerevisiae	5'-TTTTGATTCTTTGGTCACCC^TGAAGTATA 3'-AAAACTAAGAAACCAG^TGGGACTTCATAT
I-SceIII	S.cerevisiae	5'-ATTGGAGGTTTTGGTAAC^TATTTATTACC 3'-TAACCTCCAAAACC^ATTGATAAATAATGG
I-SceIV	S.cerevisiae	5'-TCTTTTCTCTTGATTA^GCCCTAATCTACG 3'-AGAAAAGAGAAC^TAATCGGGATTAGATGC

I-SceV	S.cerevisiae	5'-AATAATTTTCT^TCTTAGTAATGCC 3'-TTATTAAAAGAAGAATCATTA^CGG
I-SceVI	S.cerevisiae	5'-GTTATTTAATG^TTTTAGTAGTTGG 3'-CAATAAATTACAAAATCATCA^ACC
I-SceVII	S.cerevisiae	5'-TGTCACATTGAGGTGCACTAGTTATTAC
PI-SceI	S.cerevisiae	5'-ATCTATGTCGGGTGC^GGAGAAAGAGGTAAT 3'-TAGATACAGCC^CACGCCTCTTTCTCCATTA
F-SceI	S.cerevisiae	5'-GATGCTGTAGGC^ATAGGCTTGGTT 3'-CTACGACA^TCCGTATCCGAACCAA
F-SceII	S.cerevisiae	5'-CTTTCCGCAACA^GTAAAATT 3'-GAAAGGCG^TTGTCATTTTAA
I-HmuI	Bacillus subtilis bacteriophage SPO1	5'-AGTAATGAGCCTAACGCTCAGCAA 3'-TCATTACTCGGATTGC^GAGTCGTT
I-HmuII	Bacillus subtilis bacteriophage SP82	5'-AGTAATGAGCCTAACGCTCAACAANNNNNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNN
I-LlaI	Lactococcus lactis	5'-CACATCCATAAC^CATATCATTTTT 3'-GTGTAGGTATTGGTATAGTAA^AAA
I-MsoI	Monomastix species	5'-CTGGGTTCAAAACGTCGTGA^GACAGTTTGG 3'-GACCCAAGTTTTGCAG^CACTCTGTCAAACC
I-NanI	Naegleria andersoni	5'-AAGTCTGGTGCCA^GCACCCGC 3'-TTCAGACC^ACGGTCGTGGGCG
I-NitI	Naegleria italica	5'-AAGTCTGGTGCCA^GCACCCGC 3'-TTCAGACC^ACGGTCGTGGGCG
I-NjaI	Naegleria jamiesoni	5'-AAGTCTGGTGCCA^GCACCCGC 3'-TTCAGACC^ACGGTCGTGGGCG
I-PakI	Pseudendoclonium akinetum	5'-CTGGGTTCAAAACGTCGTGA^GACAGTTTGG 3'-GACCCAAGTTTTGCAG^CACTCTGTCAAACC
I-PorI	Pyrobaculum organotrophum	5'-GCGAGCCCGTAAGGGT^GTGTACGGG 3'-CGCTCGGGCATT^CCCACACATGCCC
I-PpoI	Physarum polycephalum	5'-TAACTATGACTCTCTTAA^GGTAGCCAAAT 3'-ATTGATACTGAGAG^AATTCCATCGGTTTA
I-ScaI	Saccharomyces capensis	5'-TGTCACATTGAGGTGCACT^AGTTATTAC 3'-ACAGTGTAACTCCAC^GTGATCAATAATG

I-Ssp6803I	Synechocystis species	5'-GTCGGGCT^CATAACCCGAA 3'-CAGCCCGAGTA^TTGGGCTT
PI-PfuI	Pyrococcus furiosus Vc1	5'-GAAGATGGGAGGAGGG^ACCGGACTCAACTT 3'-CTTCTACCCTCC^TCCCTGGCCTGAGTTGAA
PI-PfuII	Pyrococcus furiosus Vc1	5'-ACGAATCCATGTGGAGA^AGAGCCTCTATA 3'-TGCTTAGGTACAC^CTCTTCTCGGAGATAT
PI-PkoI	Pyrococcus kodakaraensis KOD1	5'-GATTTTAGAT^CCCTGTACC 3'-CTAAAA^TCTAGGGACATGG
PI-PkoII	Pyrococcus kodakaraensis KOD1	5'-CAGTACTACG^GTTAC 3'-GTCATG^ATGCCAATG
PI-PspI	Pyrococcus sp.	5'-AAAATCCTGGCAAACAGCTATTAT^GGGTAT 3'-TTTTAGGACCGTTTGTCGAT^AATACCCATA
PI-TfuI	Thermococcus fumicolans ST557	5'-TAGATTTTAGGT^CGCTATATCCTTCC 3'-ATCTAAAA^TCCAGCGATATAGGAAGG
PI-TfuII	Thermococcus fumicolans ST557	5'-TAYGCNGAYACN^GACGGYTTYT 3'-ATRCGNCT^RTGNCTGCCRAARA
PI-ThyI	Thermococcus hydrothermal-is	5'-TAYGCNGAYACN^GACGGYTTYT 3'-ATRCGNCT^RTGNCTGCCRAARA
PI-TliI	Thermococcus litoralis	5'-TAYGCNGAYACNGACGG^YTTYT 3'-ATRCGNCTRTGNC^TGCCRAARA
PI-TliII	Thermococcus litoralis	5'-AAATTGCTTGCAAACAGCTATTACGGCTAT
I-TevI	Bacteriophage T4	5'-AGTGGTATCAAC^GCTCAGTAGATG 3'-TCACCATAGT^TGCGAGTCATCTAC
I-TevII	Bacteriophage T4	5'-GCTTATGAGTATGAAGTGAACACGT^TATTC 3'-CGAATACTCATACTTCACTTGTG^CAATAAG
F-TevI	Bacteriophage T4	5'-GAAACACAAGA^AATGTTTAGTAAANNNNNNNNNNNNNN 3'-CTTTGTGTTCTTTACAAATCATTTNNNNNNNNNNNNNN^
F-TevII	Bacteriophage T4	5'-TTTAATCCTCGCTTC^AGATATGGCAACTG 3'-AAATTAGGAGCGA^AGTCTATACCGTTGAC

Relatively small deviations (degenerations) of the recognition sequence which nevertheless make possible recognition and cleavage by the particular DSBI enzyme are also included here. Such

deviations, also in connection with different basic conditions such as, for example, calcium or magnesium concentration, have been described (Argast GM et al. (1998) J Mol Biol 280:345-353). Core sequences of these recognition sequences are also included. It is known that the inner portions of the recognition sequences also suffice for an induced double-strand break and that the outer portions are not necessarily relevant but may contribute to determining the cleavage efficiency. Thus, for example, an 18bp core sequence can be defined for I-SceI.

Said DSBI recognition sequences may be localized in various positions in or close to a marker protein gene and, for example when the marker protein used is a transgene, may already be incorporated when constructing the marker protein expression cassette. Various possible localizations are illustrated by way of example in Figs. 2-A, 2-B, 3 and 5 and in the descriptions thereof.

In a further advantageous embodiment, the insertion sequence comprises at least one homology sequence A which has a sufficient length and a sufficient homology to a sequence A' in the marker protein gene in order to ensure homologous recombination between A and A'. The insertion sequence is preferably flanked by two sequences A and B which have a sufficient length and a sufficient homology to a sequence A' and, respectively, B' in the marker protein gene in order to ensure homologous recombination between A and A' and, respectively, B and B'.

"Sufficient length" means, with respect to the homology sequences A, A' and B, B', preferably sequences with a length of at least 100 base pairs, preferably at least 250 base pairs, particularly preferably at least 500 base pairs, very particularly preferably at least 1000 base pairs, most preferably of at least 2500 base pairs.

"Sufficient homology" means, with respect to the homology sequences, preferably sequences whose homology to one another is at least 70%, preferably 80%, preferentially at least 90%, particularly preferably at least 95%, very particularly preferably at least 99%, most preferably 100%, over a length of at least 20 base pairs, preferably at least 50 base pairs, particularly preferably at least 100 base pairs, very particularly preferably at least 250 base pairs, most preferably at least 500 base pairs.

Homology between two nucleic acids means the identity of the nucleic acid sequence over in each case the entire sequence length, which identity is calculated by way of comparison with the aid of

the GAP program algorithm (Wisconsin Package Version 10.0, University of Wisconsin, Genetics Computer Group (GCG), Madison, USA), setting the following parameters:

Gap Weight: 12	Length Weight: 4
Average Match: 2,912	Average Mismatch:-2,003

In a further preferred embodiment, the recombination efficiency is increased by a combination with processes which promote homologous recombination. Such systems have been described and comprise, by way of example, expression of proteins such as RecA or treatment with PARP inhibitors. It has been demonstrated that the intrachromosomal homologous recombination in tobacco plants can be increased by using PARP inhibitors (Puchta H et al. (1995) Plant J 7:203-210). The use of these inhibitors can further increase the rate of homologous recombination in the recombinant constructs, after inducing the sequence-specific DNA double-strand break, and thus the efficiency of the deletion of the transgene sequences. Various PARP inhibitors may be used here. Preference is given to including inhibitors such as 3-amino benzamide, 8-hydroxy-2-methylquinazolin-4-one (NU1025), 1,11b-dihydro-[2H]benzopyrano[4,3,2-de]isoquinolin-3-one (GPI 6150), 5-aminoisoquinolinone, 3,4-dihydro-5-[4-(1-piperidinyl)butoxy]-1(2H)-isoquinolinone or the substances described in WO 00/26192, WO 00/29384, WO 00/32579, WO 00/64878, WO 00/68206, WO 00/67734, WO 01/23386 and WO 01/23390.

Further suitable methods are the introduction of nonsense mutations into endogenous marker protein genes, for example by means of introducing RNA/DNA oligonucleotides into the plant (Zhu et al. (2000) Nat Biotechnol 18(5):555-558). Point mutations may also be generated by means of DNA-RNA hybrids which are also known as "chimeraplasty" (Cole-Strauss et al. (1999) Nucl Acids Res 27(5):1323-1330; Kmiec (1999) Gene therapy American Scientist 87(3):240-247).

The methods of dsRNAi, cosuppression by means of sense RNA and VIGS (virus induced gene silencing) are also referred to as post-transcriptional gene silencing (PTGS). PTGS processes are particularly advantageous because the demands on the homology between the marker protein gene to be reduced and the transgenically expressed sense or dsRNA nucleic acid sequence are lower than, for example, in the case of a traditional antisense approach. Thus it is possible, using the marker protein nucleic acid sequences from one species, to effectively reduce also expression of homologous

marker protein proteins in other species, without it being absolutely necessary to isolate and to elucidate the structure of the marker protein homologues occurring there. Considerably less labor is therefore required.

"Introduction" comprises within the scope of the invention any processes which are suitable for introducing an "anti-marker protein" compound, directly or indirectly, into a plant or a cell, compartment, tissue, organ or seeds of said plant or generating said compound there. The introduction may result in a transient presence of an "anti-marker protein" compound (for example a dsRNA or a recombinase) or else in a permanent (stable) presence.

According to the different nature of the approaches described above, the "anti-marker protein" compound may exert its function directly (for example by way of insertion into an endogenous marker protein gene). However, said function may also be exerted indirectly after transcription into an RNA (for example in antisense approaches) or after transcription and translation into a protein (for example in the case of recombinases or DSBI enzymes). The invention comprises both directly and indirectly acting "anti-marker protein" compounds.

Introducing comprises, for example, processes such as transfection, transduction or transformation.

"Anti-marker protein" compounds thus comprises, for example, also expression cassettes capable of implementing expression (i.e. transcription and, if appropriate, translation) of, for example, an MP-dsRNA, an MP-antisenseRNA, a sequence-specific recombinase or a DSBI enzyme in a plant cell.

"Expression cassette" means within the scope of the present invention generally those constructions in which a nucleic acid sequence to be expressed is functionally linked to at least one genetic control sequence, preferably a promoter sequence. Expression cassettes preferably consist of double-stranded DNA and may have a linear or circular structure.

A functional linkage means, for example, the sequential arrangement of a promoter with a nucleic acid sequence to be transcribed (for example coding for an MP-dsRNA or a DSBI enzyme) and, if appropriate, further regulatory elements such as, for example, a terminator and/or polyadenylation signals in such a way that each of the regulatory elements can fulfill its function during transcription of the nucleic acid sequence, depending on the arrange-

ment of the nucleic acid sequences. In this context, function can mean, for example, the control of expression, i.e. transcription and/or translation, of the nucleic acid sequence (e.g. coding for an MP-dsRNA or a DSBI enzyme). In this context, control comprises, for example, initiating, increasing, controlling or suppressing the expression, i.e. transcription and, if appropriate, translation. This does not necessarily require a direct linkage in the chemical sense. Genetic control sequences such as, for example, enhancer sequences, may exert their function on the target sequence also from positions further afar or even from different DNA molecules. Preference is given to arrangements in which the nucleic acid sequence to be transcribed is positioned downstream of the sequence acting as promoter so that both sequences are covalently connected to one another. The distance between the promoter sequence and the nucleic acid sequence to be expressed transgenically is here preferably less than 200 base pairs, particularly preferably less than 100 base pairs, very particularly preferably less than 50 base pairs.

The skilled worker knows various ways of obtaining any of the expression cassettes of the invention. An expression cassette of the invention is prepared, for example, preferably by direct fusion of a nucleic acid sequence acting as promoter to a nucleotide sequence to be expressed (e.g. coding for an MP-dsRNA or a DSBI enzyme). A functional linkage may be produced by means of common recombination and cloning techniques, as are described, for example, in Maniatis T, Fritsch EF and Sambrook J (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY and in Silhavy TJ et al. (1984) Experiments with Gene Fusions, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY and in Ausubel FM et al.(1987) Current Protocols in Molecular Biology, Greene Publishing Assoc. and Wiley Interscience.

The expression cassettes of the invention preferably comprise a promoter 5' upstream of the particular nucleic acid sequence to be expressed transgenically and a terminator sequence as an additional genetic control sequence 3' downstream and also, if appropriate, further customary regulatory elements, in each case functionally linked to the nucleic acid sequence to be expressed transgenically.

The term "genetic control sequences" is to be understood broadly and means all those sequences which have an influence on the making or function of the expression cassette of the invention. For example, genetic control sequences ensure transcription and, if

appropriate, translation in prokaryotic or eukaryotic organisms. Genetic control sequences are described, for example, in "Goeddel; Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, CA (1990)" or "Gruber and Crosby, in: Methods in Plant Molecular Biology and Biotechnolgy, CRC Press, Boca Raton, Florida, eds.:Glick and Thompson, Chapter 7, 89-108" and in the references quoted there.

Genetic control sequences comprise, in particular in plants, functional promoters. Preferred promoters suitable for the expression cassettes are in principle any promoters capable of controlling expression of genes, in particular foreign genes, in plants.

Plant-specific promoters or promoters functional in plants or in a plant cell means in principle any promoter capable of controlling expression of genes, in particular foreign genes, in at least one plant or one part, cell, tissue, culture of a plant. In this context, expression may be, for example, constitutive, inducible or development-dependent. Preference is given to:

a) Constitutive promoters

"Constitutive" promoters means those promoters which ensure expression in numerous, preferably all, tissues over a relatively large period of plant development, preferably at all points in time of plant development (Benfey et al.(1989) EMBO J 8:2195-2202). Preference is given in particular to using a plant promoter or a promoter which is derived from a plant virus. Particular preference is given to the promoter of the 35S transcript of the CaMV cauliflower mosaic virus (Franck et al. (1980) Cell 21:285-294; Odell et al. (1985) Nature 313:810-812; Shewmaker et al. (1985) Virology 140:281-288; Gardner et al. (1986) Plant Mol Biol 6:221- 228) or the 19S CaMV promoter (US 5,352,605; WO 84/02913; Benfey et al. (1989) EMBO J 8:2195-2202) and also to the promoter of the Arabidopsis thaliana nitrilase-1 gene (GenBank Acc. No.: Y07648, nucleotides 2456 (alternatively 2861) to 4308 or alternatively 4340 or 4344. (e.g. bp 2456 to 4340).

Another suitable constitutive promoter is the rubisco small subunit (SSU) promoter (US 4,962,028), the leguminB promoter (GenBank Acc. No.: X03677), the promoter of the Agrobacterium nopaline synthase, the TR dual promoter, the Agrobacterium OCS (octopine synthase) promoter, the ubiquitin promoter (Holtorf S et al. (1995) Plant Mol Biol 29:637-649), the ubi-

quitin 1 promoter (Christensen et al. (1992) Plant Mol Biol 18:675-689; Bruce et al. (1989) Proc Natl Acad Sci USA 86:9692-9696), the Smas promoter, the cinnamyl alcohol dehydrogenase promoter (US 5,683,439), the promoters of the vacuolar ATPase subunits or the promoter of a proline-rich protein from wheat (WO 91/13991), and further promoters of genes whose constitutive expression in plants is known to the skilled worker.

b) Tissue-specific promoters

Preference is given to promoters with specificities for the anthers, ovaries, flowers, leaves, stems, roots or seeds.

Seed-specific promoters comprise, for example, the promoter of phaseolin (US 5,504,200; Bustos MM et al. (1989) Plant Cell 1(9):839-53), of the 2S albumin (Joseffson LG et al. (1987) J Biol Chem 262:12196-12201), of legumin (Shirsat A et al. (1989) Mol Gen Genet 215(2): 326-331), of USP (unknown seed protein; Bäumlein H et al. (1991) Mol Gen Genet 225(3):459-67), of napin (US 5,608,152; Stalberg K et al. (1996) L Planta 199:515-519), of the sucrose-binding protein (WO 00/26388), of legumin B4 (LeB4; Bäumlein H et al. (1991) Mol Gen Genet 225: 121-128; Baeumlein et al. (1992) Plant Journal 2(2):233-9; Fiedler U et al. (1995) Biotechnology (NY) 13(10):1090f), of oleosin (WO 98/45461) or of Bce4 (WO 91/13980). Further suitable seed-specific promoters are those of the genes coding for the high molecular weight glutenin (HMWG), gliadin, branching enzyme, ADP glucose pyrophosphatase (AGPase) or starch synthase. Preference is further given to promoters which allow seed-specific expression in monocotyledones such as corn, barley, wheat, rye, rice, etc. promoters which may be employed advantageously are the promoter of the lpt2 or lpt1 gene (WO 95/15389, WO 95/23230) and the promoters described in WO 99/16890 (hordein, glutelin, oryzin, prolamin, gliadin, zein, kasirin or secalin promoters). Further seed-specific promoters are described in WO 89/03887.

Tuber-, storage-root- or root-specific promoters comprise, for example, the class I patatin promoter (B33) or the promoter of the potato cathepsin D inhibitor.

Leaf-specific promoters comprise, for example, the promoter of the potato cytosolic FBPase (WO 97/05900), the

SSU promoter (small subunit) of rubisco (ribulose-1,5-bisphosphate carboxylase) or the potato ST-LSI promoter (Stockhaus et al. (1989) EMBO J 8:2445-2451).

Flower-specific promoters comprise, for example, the phytoene synthase promoter (WO 92/16635) or the promoter of the P-rr gene (WO 98/22593).

- Anther-specific promoters comprise, for example, the 5126 promoter (US 5,689,049, US 5,689,051), the glob-1 promoter and the γ-zein promoter.

c) Chemically inducible promoters

Chemically inducible promoters allow expression control as a function of an exogenous stimulus (review article: Gatz et al. (1997) Ann Rev Plant Physiol Plant Mol Biol 48:89-108). Examples which may be mentioned are: the PRP1 promoter (Ward et al. (1993) Plant Mol Biol 22:361-366), a salicylic acid-inducible promoter (WO 95/19443), a benzenesulfonamide-inducible promoter (EP-A 0 388 186), a tetracycline-inducible promoter (Gatz et al. (1992) Plant J 2:397-404), an abscisic acid-inducible promoter (EP 0 335 528) and an ethanol- or cyclohexanone-inducible promoter (WO 93/21334). Also suitable is the promoter of the glutathione S-transferase isoform II gene (GST-II-27), which may be activated by exogenously applied safeners such as, for example, N,N-diallyl-2,2-dichloroacetamide (W0 93/01294) and which is functional in numerous tissues of both monocotyledones and dicotyledones.

Particular preference is given to constitutive or inducible promoters.

Preference is further given to plastid-specific promoters for targeted expression in the plastids. Suitable promoters are described, for example, in WO 98/55595 or WO 97/06250. promoters which may be mentioned here are the rpo B promoter element, the atoB promoter element, the clpP promoter element (see also WO 99/46394) and the 16SrDNA promoter element. Viral promoters are also suitable (WO 95/16783).

Targeted expression in plastids may also be achieved by using, for example, a bacterial or bacteriophage promoter, introducing the resulting expression cassette into the plastid DNA and then expressing expression by means of a fusion protein of a bacterial or bacteriophage polymerase and a plastid transit peptide. US

5,925,806 describes an appropriate process.

Genetic control sequences further comprise also the 5'-untranslated regions, introns or noncoding 3' region of genes, such as, for example, the actin-1 intron, or the Adh1-S introns 1, 2 and 6 (general overview: The Maize Handbook, Chapter 116, Freeling and Walbot, Eds., Springer, New York (1994)). These sequences have been shown to be able to play a significant functions in the regulation of gene expression. Thus it has been demonstrated that 5'-untranslated sequences may increase transient expression of heterologous genes. They may further promote tissue specificity (Rouster J et al.(1998) Plant J. 15:435-440). As an example of translation enhancers, mention may be made of the 5' leader sequence of the tobacco mosaic virus (Gallie et al. (1987) Nucl Acids Res 15:8693-8711).

Polyadenylation signals suitable as control sequences are in particular polyadenylation signals of plant genes and also *Agrobacterium tumefaciens* T-DNA polyadenylation signals. Examples of particularly suitable terminator sequences are the OCS (octopine synthase) terminator and the NOS (nopaline synthase) terminator (Depicker A et al (1982) J Mol Appl Genet 1:561-573) and also the terminators of soybean actin, RUBISCO or alpha-amylase from wheat (Baulcombe DC et al (1987) Mol Gen Genet 209:33-40).

Advantageously, the expression cassette may contain one or more "enhancer sequences" functionally linked to the promoter, which make increased transgenic expression of the nucleic acid sequence possible.

Genetic control sequences further means sequences coding for fusion proteins consisting of a signal peptide sequence. The expression of a target gene is possible in any desired cell compartment, such as, for example, the endomembrane system, the vacuole and the chloroplasts. Desired glycosylation reactions, in particular foldings, and the like are possible by utilizing the secretory pathway. Secretion of the target protein to the cell surface or secretion into the culture medium, for example when using suspension-cultured cells or protoplasts, is also possible. The target sequences required for this may both be taken into account in individual vector variations and be introduced into the vector together with the target gene to be cloned by using a suitable cloning strategy. Target sequences which may be used are both endogenous, if present, and heterologous sequences. Additional heterologous sequences which are preferred for functional linkage but not limited thereto are further targeting sequences

for ensuring subcellular localization in the apoplast, in the vacuole, in plastids, in the mitochrondrion, in the endoplasmic reticulum (ER), in the nucleus, in elaioplasts or other compartments; and also translation enhancers such as the 5' leader sequence from tobacco mosaic virus (Gallie et al. (1987) Nucl Acids Res 15: 8693-8711) and the like. The process of transporting proteins which are per se not located in the plastids specifically into said plastids has been described (Klosgen RB and Weil JH (1991) Mol Gen Genet 225(2):297-304; Van Breusegem F et al. (1998) Plant Mol Biol 38(3):491-496).

Control sequences are furthermore understood to be those which make possible a homologous recombination or insertion into the genome of a host organism or allow the removal from the genome. Methods such as the cre/lox technique allow the expression cassette to be removed tissue-specifically, possibly inducibly from the genome of the host organism (Sauer B. Methods. 1998; 14(4):381-92). Here, particular flanking sequences are attached to the target gene (lox sequences), which make subsequent removal by means of the cre recombinase possible.

Preferably, the expression cassette, consisting of a linkage of the promoter to the nucleic acid sequence to be transcribed, may have been integrated into a vector and may be transferred into the plant cell or organism, for example, by transformation, according to any of the processes described below.

"Transgenic" means preferably, for example with respect to a transgenic expression cassette, a transgenic expression vector, a transgenic organism or to processes for transgenic expression of nucleic acids, all constructions brought about by genetic engineering methods or processes using said constructions, in which either

a) the nucleic acid sequence to be expressed, or

b) the promoter functionally linked to the nucleic acid sequence to be expressed according to a), or

c) (a) and (b)

are not located in their natural, genetic environment (i.e. at their natural chromosomal locus) or have been modified by genetic engineering methods, the modification possibly being, for example, a substitution, addition, deletion, inversion or insertion of one or more nucleotide residues. Natural genetic environment

means the natural chromosomal locus in the source organism or the presence in a genomic library.

"Transgenic" means, with respect to expression ("transgenic expression"), preferably all expressions achieved using a transgenic expression cassette, transgenic expression vector or transgenic organism, according to the definitions indicated above.

The DNA constructs employed within the scope of the process of the invention and the vectors derived therefrom may contain further functional elements. The term functional element is to be understood broadly and means all of those elements which influence the preparation, propagation or function of the DNA constructs or of vectors or organisms derived therefrom. Examples which may be mentioned without being limited thereto are:

1. Selection markers

Selection markers comprise, for example, those nucleic acid or protein sequences whose expression gives to a cell, tissue or organism an advantage (positive selection marker) or disadvantage (negative selection marker) over cells which do not express said nucleic acid or protein. Positive selection markers act, for example, by detoxifying a substance acting on the cell in an inhibitory manner (e.g. resistance to antibiotics/herbicides) or by forming a substance which enables the plant to regenerate better or grow more under the chosen conditions (for example nutritive markers, hormone-producing markers such as ipt; see below). Another type of positive selection mareker comprises mutated proteins or RNAs which are not sensitive to a selective agent (e.g. 16S rRNA mutants which are insensitive to spectinomycin). Negative selection markers act, for example, by catalyzing the formation of a toxic substance in the transformed cells (e.g. the codA gene).

1.1 Positive selection markers:

In order to further increase the efficiency, the DNA constructs may comprise additional positive selection markers. In a preferred embodiment, the process of the invention may thus be carried out in the form of a dual selection in which a sequence coding for a resistance to at least one toxin, antibiotic or herbicide is introduced together with the nucleic acid sequence to be inserted and selection is carried out additionally by using the toxin, antibiotic or herbicide.

Appropriate proteins and sequences of positive selection markers and also selection processes are familiar to the skilled worker. The selection marker imparts to the successfully transformed cells a resistance to a biocide (e.g. a herbicide such as phosphinothricin, glyphosate or bromoxynil), a metabolism inhibitor such as 2-deoxyglucose 6-phosphate (WO 98/45456) or an antibiotic such as, for example, tetracycline, ampicillin, kanamycin, G 418, neomycin, bleomycin or hygromycin. Selection markers which may be mentioned by way of example are:

- phosphinothricin acetyltransferases (PAT) which acetylate the free amino group of the glutamine synthase inhibitor phosphinothricin (PPT) and thus detoxify PPT (de Block et al. (1987) EMBO J 6:2513-2518) (also referred to as Bialophos® resistance gene (bar)). Corresponding sequences are known to the skilled worker (from Streptomyces hygroscopicus GenBank Acc. No.: X17220 and X05822, from Streptomyces viridochromogenes GenBank Acc. No.: M 22827 and X65195; US 5,489,520). Furthermore, synthetic genes have been described for expression in plastids. A synthetic PAT gene is described in Becker et al. (1994) Plant J 5:299-307. The genes impart a resistance to the herbicide Bialaphos or glufosinate and are frequently used markers in transgenic plants (Vickers JE et al. (1996) Plant Mol Miol Reporter 14:363–368; Thompson CJ et al. (1987) EMBO J 6:2519–2523).

- 5-enolpyruvylshikimate 3-phosphate synthases (EPSPS) which impart a resistance to glyphosate (N-(phosphonomethyl) glycine). The molecular target of the unselective herbicide glyphosate is 5-enolpyruvyl-3-phosphoshikimate synthase (EPSPS). This enzyme has a key function in the biosynthesis of aromatic amino acids in microbes and plants but not in mammals (Steinrucken HC et al. (1980) Biochem Biophys Res Commun 94:1207–1212; Levin JG and. Sprinson DB (1964) J Biol Chem 239:1142–1150; Cole DJ (1985) Mode of action of glyphosate a literature analysis, p. 48–74. In: Grossbard E and Atkinson D (eds.). The herbicide glyphosate. Buttersworths, Boston.). Preference is given to using glyphosate-tolerant EPSPS variants as selection markers (Padgette SR et al. (1996). New weed control opportunities: development of soybeans with a Roundup Ready™ gene. In: Herbicide Resistant Crops (Duke, S.O., ed.), pp. 53-84. CRC Press, Boca Raton, FL; Saroha MK and Malik VS (1998) J Plant Biochemistry and Biotechnology 7:65–72). The EPSPS gene of Agrobacterium sp. strain CP4 has a natural tolerance for glyphosate, which can be transferred to appropriate transgenic plants. The CP4 EPSPS gene was cloned from Agrobacterium sp. strain CP4 (Pad-

gette SR et al. (1995)Crop Science 35(5):1451–1461). Sequences of EPSPS enzymes which are glyphosate-tolerant have been described (inter alia in US 5,510,471; US 5,776,760; US 5,864,425; US 5,633,435; US 5,627;061; US 5,463,175; EP 0 218 571). Further sequences are described under GenBank Acc. No: X63374 or M10947.

- Glyphosat®-degrading enzymes (gox gene; glyphosate oxidoreductase). GOX (for example Achromobacter sp. glyphosate oxidoreductase) catalyzes the cleavage of a C-N bond in glyphosate which is thus converted to aminomethylphosphonic acid (AMPA) and glyoxylate. GOX can thereby impart a resistance to glyphosate (Padgette SR et al. (1996) J Nutr 126(3):702-16; Shah D et al. (1986) Science 233:478-481).

- The deh gene encodes a dehalogenase which inactivates Dalapon® (GenBank Acc. No.: AX022822, AX022820 and WO 99/27116)

- The bxn genes encode bromoxynil-degrading nitrilase enzymes (Genbank Acc. No: E01313 and J03196).

- Neomycin phosphotransferases impart a resistance to antibiotics (aminoglycosides) such as neomycin, G418, hygromycin, paromomycin or kanamycin by reducing the inhibiting action of said antibiotics by means of a phosphorylation reaction. Particular preference is given to the nptII gene. Sequences can be obtained from GenBank (AF080390; AF080389). Moreover, the gene is already part of numerous expression vectors and can be isolated therefrom using processes familiar to the skilled worker (AF234316; AF234315; AF234314). The NPTII gene encodes an aminoglycoside 3'-O-phosphotransferase from E.coli, Tn5 (GenBank Acc. No: U00004 position 1401-2300; Beck et al. (1982) Gene 19 327-336).

- The DOG^R1 gene was isolated from the yeast Saccharomyces cerevisiae (EP-A 0 807 836) and encodes a 2-deoxyglucose 6-phosphate phosphatase which imparts a resistance to 2-DOG (Randez-Gil et al. (1995) Yeast 11:1233-1240; Sanz et al. (1994) Yeast 10:1195-1202, GenBank Acc. No.: NC001140; position 194799-194056).

- Acetolactate synthases which impart a resistance to imidazolinone/sulfonylurea herbicides (GenBank Acc. No.: X51514; Sathasivan K et al. (1990) Nucleic Acids Res. 18(8):2188); AB049823; AF094326; X07645; X07644; A19547; A19546; A19545;

I05376; I05373; AL133315)

- Hygromycin phosphotransferases (e.g. GenBank Acc. No.: X74325) which impart a resistance to the antibiotic hygromycin. The gene is part of numerous expression vectors and may be isolated therefrom using processes familiar to the skilled worker (such as, for example, polymerase chain reaction) (GenBank Acc. No.: AF294981; AF234301; AF234300; AF234299; AF234298; AF354046; AF354045).

- Genes of resistance to

  a) Chloramphenicol (chloramphenicol acetyltransferase),

  b) tetracycline (inter alia GenBank Acc. No.: X65876; X51366). Moreover, the gene is already part of numerous expression vectors and may be isolated therefrom using processes familiar to the skilled worker (such as, for example, polymerase chain reaction)

  c) Streptomycin (inter alia GenBank Acc. No.: AJ278607).

  d) Zeocin, the corresponding resistance gene is part of numerous cloning vectors (e.g. GenBank Acc. No.: L36849) and may be isolated therefrom using processes familiar to the skilled worker (such as, for example, polymerase chain reaction).

  e) Ampicillin (ß-lactamase gene; Datta N, Richmond MH (1966) Biochem J 98(1):204-9; Heffron F et al (1975) J. Bacteriol 122: 250–256; Bolivar F et al. (1977) Gene 2:95–114). The sequence is part of numerous cloning vectors and may be isolated therefrom using processes familiar to the skilled worker (such as, for example, polymerase chain reaction).

Genes such as isopentenyl transferase from Agrobacterium tumefaciens (strain:PO22) (Genbank Acc. No.: AB025109) may also be used as selection markers. The ipt gene is a key enzyme of cytokinin biosynthesis. Its overexpression facilitates the regeneration of plants (e.g. selection on cytokinin-free medium). The process for utilizing the ipt gene has been described (Ebinuma H et al. (2000) Proc Natl Acad Sci USA 94:2117-2121; Ebinuma H et al. (2000) Selection of Marker-free transgenic plants using the oncogenes (*ipt, rol* A, B, C) of Agrobacterium as selectable markers, In Molecular Biology of Woody Plants. Kluwer Academic Publish-

ers).

Various other positive selection markers which impart to the transformed plants a growth advantage over untransformed plants and also processes for their use are described, inter alia, in EP-A 0 601 092. Examples which may be mentioned are β-glucuronidase (in connection with cytokinin glucuronide, for example), mannose 6-phosphate isomerase (in connection with mannose), UDP-galactose 4-epimerase (in connection with galactose, for example).

For a selection marker functional in plastids, particular preference is given to those which impart a resistance to spectinomycin, streptomycin, kanamycin, lincomycin, gentamycin, hygromycin, methotrexat, bleomycin, phleomycin, blasticidin, sulfonamide, phosphinothricin, chlorsulfuron, bromoxymil, glyphosate, 2,4-datrazine, 4-methyltryptophan, nitrate, S-aminoethyl-L-cysteine, lysine/threonine, aminoethyl-cysteine or betainealdehyde. Particular preference is given to the genes aadA, nptII, BADH, FLARE-S (a fusion of aadA and GFP, described in Khan MS & Maliga P (1999) Nature Biotech 17:910-915). Especially suitable is the aadA gene (Svab Z and Maliga P (1993) Proc Natl Acad Sci USA 90:913-917). Modified 16S rDNA and also betainealdehyde dehydrogenase (BADH) from spinach have also been described (Daniell H et al. (2001) Trends Plant Science 6:237-239; Daniell H et al. (2001) Curr Genet 39:109-116; WO 01/64023; WO 01/64024; WO 01/64850). Lethal agents such as, for example, glyphosate may also be utilized in connection with correspondingly detoxifying or resistance enzymes (WO 01/81605).

The concentrations of the antibiotics, herbicides, biocides or toxins, which are used in each case for selection, must be adapted to the particular test conditions or organisms. Examples which may be mentioned for plants are kanamycin (Km) 50 mg/L, hygromycin B 40 mg/L, phosphinothricin (Ppt) 6 mg/L, spectinomycin (Spec) 500 mg/L.

2. Reporter genes

Reporter genes code for readily quantifiable proteins and thus ensure, via intrinsic color or enzyme activity, an evaluation of the transformation efficiency and of the location or time of expression. In this context, very particular preference is given to genes coding for reporter proteins (see also Schenborn E, Groskreutz D (1999) Mol Biotechnol 13(1):29-44) such as

- green fluorescence protein (GFP) (Chui WL et al. (1996) Curr Biol 6:325-330; Leffel SM et al. (1997) Biotechniques 23(5):912-8; Sheen et al. (1995) Plant J 8(5):777-784; Haseloff et al. (1997) Proc Natl Acad Sci USA 94(6): 2122-2127; Reichel et al. (1996) Proc Natl Acad Sci USA 93(12):5888-5893; Tian et al. (1997) Plant Cell Rep 16:267-271; WO 97/41228)

- chloramphenicol transferase

- luciferase (Millar et al. (1992) Plant Mol Biol Rep 10: 324-414; Ow et al. (1986) Science 234:856-859); allows bioluminescence detection

- β–galactosidase (encodes an enzyme for which various chromogenic substrates are available)

- ß-glucuronidase (GUS) (Jefferson et al. (1987) EMBO J 6: 3901-3907) or the uidA gene (encode enzymes for which various chromogenic substrates are available)

- R-locus gene product which regulates production of anthocyanin pigments (red color) in plant tissue and thus makes possible a direct analysis of the promoter activity without addition of additional auxiliary substances or chromogenic substrates (Dellaporta et al. (1988) In: Chromosome Structure and Function: Impact of New Concepts, 18th Stadler Genetics Symposium, 11:263-282)

- tyrosinase (Katz et al.(1983) J Gen Microbiol 129:2703-2714), enzyme which oxidizes tyrosine to give DOPA and dopaquinone which consequently form the readily detectable melanine.

- aequorin (Prasher et al.(1985) Biochem Biophys Res Commun 126(3):1259-1268), may be used in calcium-sensitive bioluminescence detection.

3. Origins of replication which ensure propagation of the expression cassettes or vectors of the invention, for example in E. coli. Examples which may be mentioned are ORI (origin of DNA replication), the pBR322 ori or the P15A ori (Sambrook et al.: Molecular Cloning. A Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

4. Elements, for example border sequences, which enable agrobacteria-mediated transfer into plant cells for transfer and integration into the plant genome, such as, for example, the right or left border of T-DNA or the vir region.

5. Multiple cloning regions (MCS) allow and facilitate the insertion of one or more nucleic acid sequences.

Nucleic acid sequences (e.g. expression cassettes) may be introduced into a plant organism or cells, tissues, organs, parts or seeds thereof by advantageously using vectors which contain said sequences. Vectors may be, by way of example, plasmids, cosmids, phages, viruses or else agrobacteria. The sequences may be inserted into the vector (preferably a plasmid vector) via suitable restriction cleavage sites. The resulting vector may first be introduced into E. coli and amplified. Correctly transformed E. coli are selected, grown and the recombinant vector is obtained using methods familiar to the skilled worker. Restriction analysis and sequencing may serve to check the cloning step. Preference is given to those vectors which make possible a stable integration into the host genome.

The preparation of a transformed organism (or a transformed cell or tissue) requires that the corresponding DNA (e.g. the transformation vector) or RNA is introduced into the corresponding host cell. For this process which is referred to as transformation (or transduction or transfection), a multiplicity of methods and vectors are available (Keown et al. (1990) Methods in Enzymology 185:527-537; Plant Molecular Biology and Biotechnology (CRC Press, Boca Raton, Florida), Chapter 6/7, pp. 71-119 (1993); White FF (1993) Vectors for Gene Transfer in Higher Plants; in: Transgenic Plants, Vol. 1, Engineering and Utilization, Editors: Kung and Wu R, Academic Press, 15-38; Jenes B et al. (1993) Techniques for Gene Transfer, in: Transgenic Plants, Vol. 1, Engineering and Utilization, editors: Kung and R. Wu, Academic Press, pp.128-143; Potrykus (1991) Annu Rev Plant Physiol Plant Molec Biol 42:205-225; Halford NG, Shewry PR (2000) Br Med Bull 56(1):62-73).

For example, the DNA or RNA may be introduced directly by microinjection (WO 92/09696, WO 94/00583, EP-A 0 331 083, EP-A 0 175 966) or by bombardment with DNA or RNA-coded microparticles (biolistic processes using the gene gun "particle bombardment"; US 5,100,792; EP-A 0 444 882; EP-A 0 434 616; Fromm ME et al. (1990) Bio/Technology 8(9):833-9; Gordon-Kamm et al. (1990) Plant Cell 2:603). The cell may also be permeabilized chemically, for

example with polyethylene glycol, so as to enable the DNA to reach the cell by means of diffusion. The DNA may also take place by means of protoplast fusion to other DNA-containing units such as minicells, cells, lysosomes or liposomes (Freeman et al. (1984) Plant Cell Physiol. 29:1353ff; US 4,536,475). Electroporation is another suitable method for introducing DNA, in which the cells are permeabilized reversibly by an electric impulse (EP-A 290 395, WO 87/06614). Further processes comprise the calcium-phosphate-mediated transformation, DEAE-dextran-mediated transformation, the incubation of dry embryos in DNA-containing solution or other methods of direct introduction of DNA (DE 4 005 152, WO 90/12096, US 4,684,611). Appropriate processes have been described (e.g. in Bilang et al. (1991) Gene 100:247-250; Scheid et al. (1991) Mol Gen Genet 228:104-112; Guerche et al. (1987) Plant Science 52:111-116; Neuhause et al. (1987) Theor Appl Genet 75:30-36; Klein et al. (1987) Nature 327:70-73; Howell et al. (1980) Science 208:1265; Horsch et al.(1985) Science 227:1229-1231; DeBlock et al. (1989) Plant Physiology 91:694-701; Methods for Plant Molecular Biology (Weissbach and Weissbach, eds.) Academic Press Inc. (1988); and Methods in Plant Molecular Biology (Schuler and Zielinski, eds.) Academic Press Inc. (1989)). Physical methods of introducing DNA into plant cells have been reviewed by Oard (1991) Biotech Adv 9:1-11.

In the case of these "direct" transformation methods, no particular requirements are made on the plasmid used. It is possible to use simple plasmids such as those of the pUC series, pBR322, M13mp series, pACYC184 etc.

Besides these "direct" transformation techniques, transformation may also be carried out by bacterial infection by means of Agrobacterium (e.g. EP 0 116 718), viral infection by means of viral vectors (EP 0 067 553; US 4,407,956; WO 95/34668; WO 93/03161) or by means of pollen (EP 0 270 356; WO 85/01856; US 4,684,611).

Transformation is preferably carried out by means of agrobacteria which contain disarmed Ti-plasmid vectors, using the latters' natural ability to transfer genes to plants (EP-A 0 270 355; EP-A 0 116 718). Agrobacterium transformation is widespread for transforming dicotyledones, but is also increasingly applied to monocotyledones (Toriyama et al. (1988) Bio/Technology 6: 1072-1074; Zhang et al. (1988) Plant Cell Rep 7:379-384; Zhang et al. (1988) Theor Appl Genet 76:835-840; Shimamoto et al. (1989) Nature 338:274-276; Datta et al. (1990) Bio/Technology 8: 736-740; Christou et al. (1991) Bio/Technology 9:957-962; Peng et al. (1991) International Rice Research Institute, Manila, Philippines

563-574; Cao et al. (1992) Plant Cell Rep 11:585-591; Li et al. (1993) Plant Cell Rep 12:250-255; Rathore et al. (1993) Plant Mol Biol 21:871-884; Fromm et al. (1990) Bio/Technology 8:833-839; Gordon-Kamm et al. (1990) Plant Cell 2:603-618; D'Halluin et al. (1992) Plant Cell 4:1495-1505; Walters et al. (1992) Plant Mol Biol 18:189-200; Koziel et al. (1993) Biotechnology 11:194-200; Vasil IK (1994) Plant Mol Biol 25:925-937; Weeks et al. (1993) Plant Physiol 102:1077-1084; Somers et al. (1992) Bio/Technology 10:1589-1594; WO 92/14828; Hiei et al. (1994) Plant J 6:271-282).

The strains most often used for agrobacterial transformation, Agrobacterium tumefaciens or Agrobacterium rhizogenes, contain a plasmid (Ti and Ri plasmids, respectively), which is transferred to the plant after agrobacterial infection. Part of this plasmid, called T-DNA (transferred DNA), is integrated into the genome of the plant cell. Alternatively, Agrobacterium may also transfer binary vectors (mini Ti plasmids) to plants and integrate them into the genome of said plants.

The application of Agrobacterium tumefaciens to the transformation of plants, using tissue culture explants, has been described (inter alia, Horsch RB et al. (1985) Science 225:1229ff; Fraley et al. (1983) Proc Natl Acad Sci USA 80: 4803-4807; Bevans et al. (1983) Nature 304:184-187). Many Agrobacterium tumefaciens strains are capable of transferring genetic material, such as, for example, the strains EHA101[pEHA101], EHA105[pEHA105], LBA4404[pAL4404], C58C1[pMP90] and C58C1[pGV2260] (Hood et al. (1993) Transgenic Res 2:208-218; Hoekema et al. (1983) Nature 303:179-181; Koncz and Schell (1986) Gen Genet 204:383-396; Deblaere et al. (1985) Nucl Acids Res 13: 4777-4788).

When using agrobacteria, the expression cassette must be integrated into special plasmids, either a shuttle or intermediate vector or a binary vector. When using a Ti or Ri plasmid for transformation, then at least the right border, but usually the right and left borders of the Ti or Ri plasmid T-DNA are connected as a flanking region to the expression cassette to be introduced. Preference is given to using binary vectors. Binary vectors may replicate both in E. coli and in agrobacteria and contain the components required for transfer into a plant system. They normally contain a selection marker gene for selection of transformed plants (e.g. the nptII gene which imparts a resistance to kanamycin) and a linker or polylinker flanked by the right and left T-DNA border sequences. They contain moreover, outside the T-DNA border sequence, also a selection marker which enables transformed E. coli and/or agrobacteria to be selected

(e.g. the nptIII gene which imparts a resistance to kanamycin). Corresponding vectors may be transformed directly into Agrobacterium (Holsters et al. (1978) Mol Gen Genet 163:181-187).

Binary vectors are based, for example, on "broad host range" plasmids such as pRK252 (Bevan et al. (1984) Nucl Acid Res 12,8711-8720) and pTJS75 (Watson et al. (1985) EMBO J 4(2):277-284). A large group of the binary vectors used is derived from pBIN19 (Bevan et al. (1984) Nucl Acid Res 12:8711-8720). Hajdukiewicz et al. developed a binary vector (pPZP) which is smaller and more efficient than the previously customary vectors (Hajdukiewicz et al. (1994) Plant Mol Biol 25:989-994). Improved and particularly preferred binary vector systems for Agrobacterium-mediated transformation are described in WO 02/00900.

The agrobacteria transformed with a vector of this kind may then be used in the known manner for transforming plants, in particular crop plants such as, for example, oilseed rape, for example by bathing wounded leaves or leaf sections in an agrobacterial solution and subsequently culturing them in suitable media. The transformation of plants by agrobacteria has been described (White FF, Vectors for Gene Transfer in Higher Plants; in Transgenic Plants, Vol. 1, Engineering and Utilization, edited by S.D. Kung and R. Wu, Academic Press, 1993, pp. 15-38; Jenes B et al.(1993) Techniques for Gene Transfer, in: Transgenic Plants, Vol. 1, Engineering and Utilization, edited by S.D. Kung and R. Wu, Academic Press, pp.128-143; Potrykus (1991) Annu Rev Plant Physiol Plant Molec Biol 42:205-225). Transgenic plants may be regenerated in the known manner from the transformed cells of the wounded leaves or leaf sections.

Different explants, cell plants, tissues, organs, embryos, seeds, microspores or other unicellular or multicellular cellular structures derived from a plant organism may be used for transformation. Transformation processes adjusted to the particular explants, cultures or tissues are known to the skilled worker. Examples which may be mentioned are: shoot internodes (Fry J et al. (1987) Plant Cell Rep. 6:321-325), hypocotyls (Radke SE et al. (1988) Theor Appl Genet 75:685-694; Schröder M et al. (1994) Physiologia Plant 92: 37-46.; Stefanov I et al. (1994) Plant Sci. 95:175-186; Weier et al. (1997) Fett/Lipid 99:160-165), cotyledonous petioles (Meloney MM et al. (1989) Plant Cell Rep 8:238-242; Weier D et al. (1998) Molecular Breeding 4:39-46), microspores and proembryos (Pechnan (1989) Plant Cell Rep. 8:387-390) and flower stalks (Boulter ME et al. (1990) Plant Sci 70:91-99; Guerche P et al. (1987) Mol Gen Genet 206:382-386). In the case

of a direct gene transfer, mesophyll protoplasts (Chapel PJ & Glimelius K (1990) Plant Cell Rep 9: 105-108; Golz et al. (1990) Plant Mol Biol 15:475-483) or else hypocotyl protoplasts (Bergmann P & Glimelius K (1993) Physiologia Plant 88:604-611) and microspores (Chen JL et al. (1994) Theor Appl Genet 88:187-192; Jonesvilleneuve E et al. (1995) Plant Cell Tissue and Organ Cult 40:97-100) and shoot sections (Seki M et al. (1991) Plant Mol Biol 17:259-263) may be employed successfully.

Stably transformed cells, i.e. those which contain the introduced DNA integrated into the DNA of the host cell, may be selected from untransformed cells by using the selection process of the invention. The plants obtained may be grown and crossed in the usual way. Preferably, two or more generations should be cultured in order to ensure that the genomic integration is stable and can be inherited.

As soon as a transformed plant cell has been prepared, it is possible to obtain a complete plant by using proceses known to the skilled worker. This involves, for example, starting from callus cultures, individual cells (e.g. protoplasts) or leaf disks (Vasil et al. (1984) Cell Culture and Somatic Cel Genetics of Plants, Vol I, II and III, Laboratory Procedures and Their Applications, Academic Press; Weissbach and Weissbach (1989) Methods for Plant Molecular Biology, Academic Press). It is possible to induce from these still undifferentiated callus cell masses the formation of shoot and root in the known manner. The seedlings obtained may be planted out and grown. Appropriate processes have been described (Fennell et al. (1992) Plant Cell Rep. 11: 567-570; Stoeger et al. (1995) Plant Cell Rep. 14:273-278; Jahne et al. (1994) Theor Appl Genet 89:525-533).

The efficacy of expressing the transgenically expressed nucleic acids may be determined, for example, *in vitro* by shoot-meristem propagation using any of the selection methods described above. Moreover, changes in the type and level of expression of a target gene and the effect on the phenotype of the plant may be tested in greenhouse experiments using test plants.

The process of the invention is preferably used within the framework of plant biotechnology for generating plants having advantageous properties. The "nucleic acid sequence to be inserted" into the genome of the plant cell or the plant organism preferably comprises at least one expression cassette, said expression cassette being able to express, under the control of a promoter functional in plant cells or plant organisms, an RNA and/or a

protein which do not cause reduction of the expression, amount, activity and/or function of a marker protein but, particularly preferably, impart to the plant genetically altered in this way an advantageous phenotype. Numerous genes and proteins which may be used for achieving an advantageous phenotype, for example for the increase in quality of foodstuff or for producing particular chemicals or pharmaceuticals (Dunwell JM (2000) J Exp Bot 51 Spec No:487-96) are known to the skilled worker.

Thus it is possible to improve the suitability of the plants or the seeds thereof as foodstuff or feedstuff, for exmaple by altering the compositions and/or the content of metabolites, in particular proteins, oils, vitamins and/or starch. It is also possible to increase the growth rate, yield or resistance to biotic or abiotic stress factors. Advantageous effects may be achieved both by transgenic expression of nucleic acids or proteins and by targeted reduction of the expression of endogenous genes, with respect to the phenotype of the transgenic plant. The advantageous effects which may be achieved in the transgenic plant comprise, for example:

- increased resistance to pathogens (biotic stress)

- increased resistance to environmental factors such as heat, cold, frost, drought, UV light, oxidative stress, wetness, salt, etc. (abiotic stress)

- increased yield

- improved quality, for example increased nutritional value, increased storability

The invention further relates to the use of the transgenic plants prepared according to the process of the invention and of the cells, cell cultures, plants or propagation material such as seeds or fruits derived from said plants, for preparing foodstuff or feedstuff, pharmaceuticals or fine chemicals such as, for example, enzymes, vitamins, amino acids, sugars, fatty acids, natural and synthetic flavorings, aroma substances and colorants. Particular preference is given to the production of triacyl glycerides, lipids, oils, fatty acids, starch, tocopherols and tocotrienols and also carotenoids. Genetically modified plants of the invention, which may be consumed by humans and animals may also be used as foodstuff or feedstuff, for example, directly or after preparation known per se.

As already mentioned above, the process of the invention comprises in a particularly advantageous embodiment, in a process step downstream of the selection, the deletion of the sequence coding for the marker protein (e.g. mediated by recombinase or as described in WO03/004659) or the elimination by crossing and/or segregation of said sequences. (It is obvious to the skilled worker that, for this purpose, the nucleic acid sequence integrated into the genome and the sequence coding for the marker protein should have a separate chromosomal locus in the transformed cells. This, however, is the case in the majority of the resulting plants, merely for reasons of statistics). This procedure is particularly advantageous if the marker protein is a transgene which otherwise does not occur in the plant to be transformed. Although the resulting plant may still possibly contain the compound for reducing the expression, amount, activity and/or function of the marker protein, said compound would have no longer any "counterpart" in the form of said marker protein, and thus would have no effect. This is particularly the case if the marker protein is derived from a non-plant organism and/or is synthetic (for example the codA protein). It is, however, also possible to use plant marker proteins from other plant species, which otherwise do not occur in the cell to be transformed (i.e. if not introduced as transgene). Said marker proteins are referred to as "nonendogenous" marker proteins within the scope of the present invention.

Very particularly advantageously, the compound for reducing the expression, amount, activity and/or function of the marker protein is an RNA. After deletion or elimination by crossing/segregation, the resulting transgenic plant would have no longer any unnecessary (and, if appropriate, undesired) foreign protein. The sole foreign protein would be possibly the protein resulting from the nucleic acid sequence inserted into the genome. For reasons of product approval, this embodiment is particularly advantageous. As described above, said RNA may be an antisense RNA or, particularly preferably, a double-stranded RNA. It may be expressed separately from the RNA coding for the target protein but also, possibly, on the same strand as the latter.

In summary, the particularly advantageous embodiment comprises the following features:

A process for preparing transformed plant cells or organisms, which comprises the following steps:

a) transforming a population of plant cells which comprises at least one non-endogenous (preferably non-plant) marker pro-

tein capable of converting directly or indirectly a substance X which is nontoxic for said population of plant cells into a substance Y which is toxic for said population, with at least one nucleic acid sequence to be inserted in combination with at least one nucleic acid sequence coding for a ribonucleic acid sequence capable of reducing the expression, amount, activity and/or function of said marker protein, and

b) treating said population of plant cells with the substance X at a concentration which causes a toxic effect for nontransformed cells, due to the conversion by the marker protein, and

c) selecting transformed plant cells (and/or populations of plant cells, such as plant tissues or plants) whose genome contains said nucleic acid sequence and which have a growth advantage over nontransformed cells, due to the action of said compound, from said population of plant cells, the selection being carried out under conditions under which the marker protein can exert its toxic effect on the nontransformed cells, and

d) regenerating fertile plants, and

e) eliminating by crossing the nucleic acid sequence coding for the marker protein and isolating fertile plants whose genome contains said nucleic acid sequence but does not contain any longer the sequence coding for the marker protein.

Sequences

SEQ ID NO: 1 Nucleic acid sequence coding for E. coli cytosine deaminase (codA)

SEQ ID NO: 2 amino acid sequence coding for E. coli cytosine deaminase (codA)

SEQ ID NO: 3 Nucleic acid sequence coding for E. coli cytosine deaminase (codA), with modified start codon (GTG/ATG) for expression in eukaryotes

SEQ ID NO: 4 Amino acid sequence coding for E. coli cytosine deaminase (codA), with modified start codon (GTG/ATG) for expression in eukaryotes

SEQ ID NO: 5 Nucleic acid sequence coding for Streptomyces griseolus cytochrome P450-SU1 (suaC)

SEQ ID NO: 6 Amino acid sequence coding for Streptomyces griseolus cytochrome P450-SU1 (suaC)

SEQ ID NO: 7 Nucleic acid sequence coding for Agrobacterium tumefaciens indoleacetamide hydrolase (tms2)

SEQ ID NO: 8 Amino acid sequence coding for Agrobacterium tumefaciens indoleacetamide hydrolase (tms2)

SEQ ID NO: 9 Nucleic acid sequence coding for Agrobacterium tumefaciens indoleacetamide hydrolase (tms2)

SEQ ID NO: 10 Amino acid sequence coding for Agrobacterium tumefaciens indoleacetamide hydrolase (tms2)

SEQ ID NO: 11 Nucleic acid sequence coding for Xanthobacter autotrophicus haloalkane dehalogenase (dhlA)

SEQ ID NO: 12 Amino acid sequence coding for Xanthobacter autotrophicus haloalkane dehalogenase (dhlA)

SEQ ID NO: 13 Nucleic acid sequence coding for Herpes simplex Virus 1 thymidine kinase

SEQ ID NO: 14 Amino acid sequence coding for Herpes simplex Virus 1 thymidine kinase

SEQ ID NO: 15 Nucleic acid sequence coding for Herpes simplex Virus 1 thymidine kinase

SEQ ID NO: 16 Amino acid sequence coding for Herpes simplex Virus 1 thymidine kinase

SEQ ID NO: 17 Nucleic acid sequence coding for Toxoplasma gondii hypoxanthine-xanthine-guanine phosphoribosyl transferase

SEQ ID NO: 18 Amino acid sequence coding for Toxoplasma gondii hypoxanthine-xanthine-guanine phosphoribosyl transferase

SEQ ID NO: 19 Nucleic acid sequence coding for E. coli xanthine-guanine phosphoribosyl transferase

SEQ ID NO: 20 Amino acid sequence coding for E. coli xanthine-guanine phosphoribosyl transferase

SEQ ID NO: 21 Nucleic acid sequence coding for E. coli xanthine-guanine phosphoribosyl transferase

SEQ ID NO: 22 Amino acid sequence coding for E. coli xanthine-guanine phosphoribosyl transferase

SEQ ID NO: 23 Nucleic acid sequence coding for E. coli purine nucleoside phosphorylase (deoD)

SEQ ID NO: 24 Nucleic acid sequence coding for E. coli purine nucleoside phosphorylase (deoD)

SEQ ID NO: 25 Nucleic acid sequence coding for Burkholderia caryophylli phosphonate monoester hydrolase (pehA)

SEQ ID NO: 26 Amino acid sequence coding for Burkholderia caryophylli phosphonate monoester hydrolase (pehA)

SEQ ID NO: 27 Nucleic acid sequence coding for Agrobacterium rhizogenes tryptophan oxygenase (aux1)

SEQ ID NO: 28 Amino acid sequence coding for Agrobacterium rhizogenes tryptophan oxygenase (aux1)

SEQ ID NO: 29 Nucleic acid seuence coding for Agrobacterium rhizogenes indoleacetamide hydrolase (aux2)

SEQ ID NO: 30 Amino acid seuence coding for Agrobacterium rhizogenes indoleacetamide hydrolase (aux2)

SEQ ID NO: 31 Nucleic acid sequence coding for Agrobacterium tumefaciens tryptophan oxygenase (aux1)

SEQ ID NO: 32 Amino acid sequence coding for Agrobacterium tumefaciens tryptophan oxygenase (aux1)

SEQ ID NO: 33 Nucleic acid sequence coding for Agrobacterium tumefaciens indoleacetamide hydrolase (aux2)

SEQ ID NO: 34 Amino acid sequence coding for Agrobacterium tumefaciens indoleacetamide hydrolase (aux2)

SEQ ID NO: 35 Nucleic acid sequence coding for Agrobacterium vitis indoleacetamide hydrolase (aux2)

SEQ ID NO: 36 Amino acid sequence coding for Agrobacterium vitis indoleacetamide hydrolase (aux2)

SEQ ID NO: 37 Nucleic acid sequence coding for Arabidopsis thaliana 5-methylthioribose kinase (mtrK)

SEQ ID NO: 38 Amino acid sequence coding for Arabidopsis thaliana 5-methylthioribose kinase (mtrK)

SEQ ID NO: 39 Nucleic acid sequence coding for Klebsiella pneumoniae 5-methylthioribose kinase (mtrK)

SEQ ID NO: 40 Amino acid sequence coding for Klebsiella pneumoniae 5-methylthioribose kinase (mtrK)

SEQ ID NO: 41 Nucleic acid sequence coding for Arabidopsis thaliana alcohol dehydrogenase (adh)

SEQ ID NO: 42 Amino acid sequence coding for Arabidopsis thaliana alcohol dehydrogenase (adh)

SEQ ID NO: 43 Nucleic acid sequence coding for Hordeum vulgare (barley) alcohol dehydrogenase (adh)

SEQ ID NO: 44 Amino acid sequence coding for Hordeum vulgare (barley) alcohol dehydrogenase (adh)

SEQ ID NO: 45 Nucleic acid sequence coding for Oryza sativa (rice) alcohol dehydrogenase (adh)

SEQ ID NO: 46 Amino acid sequence coding for Oryza sativa (rice) alcohol dehydrogenase (adh)

SEQ ID NO: 47 Nucleic acid sequence coding for Zea mays (corn) alcohol dehydrogenase (adh)

SEQ ID NO: 48 Amino acid sequence coding for Zea mays (corn) alcohol dehydrogenase (adh)

SEQ ID NO: 49 Nukleic acid sequence coding for a sense RNA fragment of E. coli cytosine deaminase (codARNAi-sense)

SEQ ID NO: 50 Oligonucleotide primer codA5'HindIII
5'-AAGCTTGGCTAACAGTGTCGAATAACG-3'

SEQ ID NO: 51 Oligonucleotide primer codA3'SalI
5'-GTCGACGACAAAATCCCTTCCTGAGG-3'

SEQ ID NO: 52 Nucleic acid sequence coding for an antisense RNA fragment of E. coli cytosine deaminase (codARNAi-anti)

SEQ ID NO: 53 Oligonucleotide primer codA5'EcoRI
5'-GAATTCGGCTAACAGTGTCGAATAACG-3'

SEQ ID NO: 54 Oligonucleotide primer codA3'BamHI
5'-GGATCCGACAAAATCCCTTCCTGAGG-3'

SEQ ID NO: 55 Vector construct pBluKS-nitP-STLS1-35S-T

SEQ ID NO: 56 Expression vector pSUN-1

SEQ ID NO: 57 Transgenic expression vector pSUN-1-codA-RNAi

SEQ ID NO: 58 Transgenic expression vector pSUN1-codA-RNAi-At.Act.-2-At.Als-R-ocsT

SEQ ID NO: 59 Nukleic acid sequence coding for 5-methylthioribose kinase (mtrK) from corn (Zea mays); fragment

SEQ ID NO: 60 Amino acid sequence coding for 5-methylthioribose kinase (mtrK) from corn (Zea mays); fragment

SEQ ID NO: 61 Nucleic acid sequence coding for 5-methylthioribose kinase (mtrK) from oilseed rape (Brassica napus), fragment

SEQ ID NO: 62 Amino acid sequence coding for 5-methylthioribose kinase (mtrK) from oilseed rape (Brassica napus), fragment

SEQ ID NO: 63 Nucleic acid sequence coding for 5-methylthioribose kinase (mtrK) from oilseed rape (Brassica na-

pus), fragment

SEQ ID NO: 64 Amino acid sequence coding for 5-methylthioribose kinase (mtrK) from oilseed rape (Brassica napus), fragment

SEQ ID NO: 65 Nucleic acid sequence coding for 5-methylthioribose kinase (mtrK) from rice (Oryza sativa), fragment

SEQ ID NO: 66 Amino acid sequence coding for 5-methylthioribose kinase (mtrK) from rice (Oryza sativa), fragment

SEQ ID NO: 67 Nucleic acid sequence coding for 5-methylthioribose kinase (mtrK) from soybean (Glycine max), fragment

SEQ ID NO: 68 Amino acid sequence coding for 5-methylthioribose kinase (mtrK) from soybean (Glycine max), fragment

SEQ ID NO: 69 Oligonucleotide primer codA5'C-term
5'-CGTGAATACGGCGTGGAGTCG-3'

SEQ ID NO: 70 Oligonucleotide primer codA3'C-term
5'-CGGCAGGATAATCAGGTTGG-3'

SEQ ID NO: 71 Oligonucleotide primer 35sT 5' primer
5'-GTCAACGTAACCAACCCTGC-3'

Figures

Fig.1: Inactivation of the marker protein gene by means of introducing a recombinase

P:	promoter
MP:	Sequence coding for a marker protein
R1/R2:	Recombinase recognition sequences
R:	Recombinase or sequence coding for recombinase.

In a preferred embodiment, the marker protein gene is inactivated by introducing a sequence-specific recombinase. Preference is given to its expressing the recombinase, as depicted here, starting from an expression cassette.

The marker protein gene is flanked by recognition sequences for sequence-specific recombinases, with sequences of said marker protein gene being deleted by introducing said recombinase and thus said marker protein gene being inactivated.

Fig.2-A: Inactivation of the marker protein gene by the action of a sequence-specific nuclease

P:	promoter
DS:	Recognition sequence for targeted induction of DNA double-strand breaks
MP-DS-MP':	Sequence coding for a marker protein, comprising a DS
nDS:	Inactivated DS
E:	Sequence-specific enzyme for targeted induction of DNA double-strand breaks

The marker protein gene may be established by a targeted mutation or deletion in the marker protein gene, for example by sequence-specific induction of DNA double-strand breaks at a recognition sequence for targeted induction of DNA double-strand breaks in or close to the marker protein gene (P-MP). The double-strand break may occur in the coding region or else the noncoding (such as, for example, the promoter) region, induces an illegitimate recombination (nonhomologous DNA-end joining) and thus, for example, a shift in the reading frame of said marker protein.

Fig.2-B: Inactivation of the marker protein gene by the action of a sequence-specific nuclease

P:	promoter
DS:	Recognition sequence for targeted induction of DNA double-strand breaks
MP:	Sequence coding for a marker protein
nDS:	Inactivated DS
E:	Sequence-specific enzyme for targeted induction of DNA double-strand breaks

The marker protein gene may be established by a targeted deletion by sequence-specific induction of more than one sequence-specific DNA double-strand break in or close to said marker protein gene. The double-strand breaks may occur in the coding region or else the noncoding (such as, for example, the promoter) region and induce a deletion in the marker protein gene. The marker protein gene is preferably flanked by DS sequences and is completely deleted by the action of enzyme E.

Fig. 3: Inactivation of the marker protein gene by inducing an intramolecular homologous recombination, due to the action of a sequence-specific nuclease

A/A':	Sequences with a sufficient length and homology to one another, in order to recombine with one another as a consequence of the induced double-strand break
P:	promoter
DS:	Recognition sequence for targeted induction of DNA double-strand breaks
MP:	Sequence coding for a marker protein
E:	Sequence-specific enzyme for targeted induction of DNA double-strand breaks

The marker protein gene may be inactivated by a deletion by means of intramolecular homologous recombination. Said homologous recombination may be initiated by sequence-specific induction of DNA double-strand breaks at a recognition sequence for targeted induction of DNA double-strand breaks in or close to the marker protein gene. The homologous recombination occurs between the sequences A and A' which have a sufficient length and homology to one another in order to recombine with one another as a consequence of the induced double-strand break. The recom-

bination causes a deletion of essential sequences of the marker protein gene.

Fig. 4: Inactivation of the marker protein gene by intermolecular homologous recombination

A/A':	Sequences with a sufficient length and homology to one another in order to recombine with one another
B/B':	Sequences with a sufficient length and homology to one another in order to recombine with one another
P:	promoter
I:	nucleic acid sequence/gene of interest to be inserted
MP:	Sequence coding for a marker protein

The marker protein gene (P-MP) may also be inactivated by a targeted insertion into the marker protein gene, for example by means of intermolecular homologous recombination. In this context, the region to be inserted is flanked on its 5' and 3' ends by nucleic acid sequences (A' and B', respectively), which have a sufficient length and homology to corresponding flanking sequences of the marker protein (A and B, respectively) in order to make possible a homologous recombination between A and A' and B and B'. The recombination causes a deletion of essential sequences of the marker protein gene.

Fig. 5: Inactivation of the marker protein gene by intermolecular homologous recombination due to the action of a sequence-specific nuclease

A/A':	Sequences with a sufficient length and homology to one another in order to recombine with one another
B/B':	Sequences with a sufficient length and homology to one another in order to recombine with one another
P:	promoter
I:	nucleic acid sequence/gene of interest to be inserted
MP:	Sequence coding for a marker protein
DS:	Recognition sequence for targeted induction of DNA double-strand breaks
E:	Sequence-specific enzyme for targeted

induction of DNA double-strand breaks

The marker protein gene may also be inactivated by a targeted insertion into the marker protein gene, for example by means of intermolecular homologous recombination. The homologous recombination may be initiated by sequence-specific induction of DNA double-strand breaks at a recognition sequence for targeted induction of DNA double-strand breaks in or close to the marker protein gene. In this context, the region to be inserted is flanked at its 5' and 3' ends by nucleic acid sequences (A' and B', respectively) which have a sufficient length and homology to corresponding flanking sequences of the marker protein gene (A and B, respectively) in order to make possible a homologous recombination between A and A' and B and B'. The recombination causes a deletion of essential sequences of the marker protein gene.

Fig. 6: Vector map for pBluKS-nitP-STLS1-35S-T (SEQ ID NO: 55)

NitP: promoter of the A. thaliana nitrilaseI gene (GenBank Acc. No.: Y07648.2, Hillebrand et al. (1996) Gene 170:197-200)

STLS-1 intron: intron of the potato ST-LS1 gene (Vanconneyt GF et al. (1990) Mol Gen Genet 220(2):245-250).

35S-Term: Terminator of the 35S CaMV gene (cauliflower mosaic virus; Franck et al. (1980) Cell 21:285-294).

Cleavage sites of relevant restriction endonucleases are indicated with their particular cleavage position.

Fig. 7: Vector map for the transgenic expression vector pSUN-1-codA-RNAi (SEQ ID NO: 57)

NitP: promoter of the A. thaliana nitrilaseI gene (GenBank Acc. No.: Y07648.2, Hillebrand et al. (1996) Gene 170:197-200)

STLS-1 intron: intron of the potato ST-LS1 gene (Vanconneyt GF et al. (1990) Mol Gen Genet 220(2):245-250).

35S-Term: Terminator of the 35S CaMV gene (cauliflower mosaic virus; Franck et al. (1980) Cell 21:285-294).

codA-sense: Nucleic acid sequence coding for a sense RNA fragment of E. coli cytosine deaminase (codARNAi-sense; SEQ ID NO: 49)

codA-anti: Nucleic acid sequence coding for an antisense RNA fragment of E. coli cytosine deaminase (codARNAi-anti; SEQ ID NO: 52)

LB/RB: Left and, respectively, right boundaries of Agrobacterium T-DNA

Cleavage sites of relevant restriction endonucleases are indicated with their particular cleavage position. Further elements represent customary elements of a binary Agrobacterium vector (aadA; ColE1; repA)

Fig. 8: Vector map for the transgenic expression vector pSUN1-codA-RNAi-At.Act.-2-At.Als-R-ocsT (SEQ ID NO: 58)

NitP: promoter of the A. thaliana nitrilaseI gene (GenBank Acc. No.: Y07648.2, Hillebrand et al. (1996) Gene 170:197-200)

STLS-1 intron: intron of the potato ST-LS1 gene (Vancanneyt GF et al. (1990) Mol Gen Genet 220(2):245-250).

35S-Term: Terminator of the 35S CaMV gene (cauliflower mosaic virus; Franck et al. (1980) Cell 21:285-294).

codA-sense: Nucleic acid sequence coding for a sense RNA fragment of E. coli cytosine deaminase (codARNAi-sense; SEQ ID NO: 49)

codA-anti: Nucleic acid sequence coding for an antisense RNA fragment of E. coli cytosine deaminase (codARNAi-anti; SEQ ID NO: 52)

Left border/right border: Left and, respectively, right boundaries of Agrobacterium T-DNA

Cleavage sites of relevant restriction endonucleases are indicated with their particular cleavage position. Further elements represent customary elements of a binary Agrobacterium vector (aadA; ColE1; repA)

Fig.9a-b: Sequence comparison of various 5-methylthioribose (MTR) kinases from various organisms, in particular plant organisms. Sequences from Klebsiella pneumoniae, Clostridium tetani, Arabidopsis thaliana (A.thaliana), oilseed rape (Brassica napus), soybean (Soy-1), rice (Oryza sativa-1) and also the consensus sequence (Consensus) are shown. Homologous regions can be readily deduced from the consensus sequence.

Exemplary embodiments

General methods

The chemical synthesis of oligonucleotides may be carried out, for example, in the known manner by using the phosphoamide method (Voet, Voet, 2nd Edition, Wiley Press New York, pages 896-897). The cloning steps carried out within the scope of the present invention, such as, for example, restriction cleavages, agarose gel electrophoresis, purification of DNA fragments, transfer of nucleic acids to nitrocellulose and nylon membranes, linking of DNA fragments, transformation of E. coli cells, cultivation of bacteria, propagation of phages and sequence analysis of recombinant DNA, are carried out as described in Sambrook et al. (1989) Cold Spring Harbor Laboratory Press; ISBN 0-87969-309-6. The sequencing of recombinant DNA molecules was carried out using a laser fluorescence DNA sequencer from ABI, according to the method of Sanger (Sanger et al. (1977) Proc Natl Acad Sci USA 74:5463-5467).

Example 1: Preparation of codA fragments

First, a truncated nucleic acid variant of the *codA* gene, modified by the addition of recognition sequences of the restriction enzymes HindIII and SalI, is prepared using the PCR technique. For this purpose, part of the *codA* gene (GeneBank Acc. No.: S56903; SEQ ID NO: 1) is amplified from the E. coli source organism by means of the polymerase chain reaction (PCR) using a sense-specific primer (codA5'HindIII; SEQ ID NO: 50) and an antisense-specific primer (codA3'SalI; SEQ ID NO: 51).

codA5'HindIII: 5'-AAGCTTGGCTAACAGTGTCGAATAACG-3' (SEQ ID NO: 50)

codA3'SalI: 5'-GTCGACGACAAAATCCCTTCCTGAGG-3' (SEQ ID NO: 51)

The PCR was carried out in 50 µl reaction mixture which contained:

- 2 µl (200 ng) of *E. coli* genomic DNA
- 0.2 mM dATP, dTTP, dGTP, dCTP
- 1.5 mM $Mg(OAc)_2$
- 5 µg of bovine serum albumin
- 40 pmol of "codA5'HindIII" primer
- 40 pmol of "codA3'SalI" primer
- 15 µl of 3.3× rTth DNA Polymerase XLPuffer (PE Applied Biosystems)

- 5U of rTth DNA Polymerase XL (PE Applied Biosystems)

The PCR is carried out under the following cycle conditions:

Step 1: 5 minutes 94°C (denaturation)
Step 2: 3 seconds 94°C
Step 3: 1 minute 60°C (annealing)
Step 4: 2 minutes 72°C (elongation)

30 repeats of steps 2 to 4

Step 5: 10 minutes 72°C (post elongation)
Step 6: 4°C (waiting loop)

The amplicon (codARNAi-sense; SEQ ID NO: 49) is cloned using standard methods into the PCR cloning vector pGEM-T (Promega). The identity of the amplicon generated is confirmed by sequencing using the M13F (-40) primer.

Another truncated fragment of the *codA* gene, modified by the addition of recognition sequences of the restriction enzymes EcoRI and BamHI, is amplified using a sense-specific primer (codA5'EcoRI; SEQ ID NO: 53) and an antisense-specific primer (codA3'BamHI; SEQ ID NO: 54).

codA5'EcoRI: 5'-GAATTCGGCTAACAGTGTCGAATAACG-3' (SEQ ID NO: 53)

codA3'BamHI: 5'-GGATCCGACAAAATCCCTTCCTGAGG-3' (SEQ ID NO: 54)

The PCR was carried out in 50 µl reaction mixture which contained:

- 2 µl (200 ng) of *E. coli* genomic DNA
- 0.2 mM dATP, dTTP, dGTP, dCTP
- 1.5 mM $Mg(OAc)_2$
- 5 µg of bovine serum albumin
- 40 pmol of "codA5'EcoRI" primer
- 40 pmol of "codA3'BamHI" primer
- 15 µl of 3.3x rTth DNA Polymerase XLPuffer (PE Applied Biosystems)
- 5U of rTth DNA Polymerase XL (PE Applied Biosystems)

The PCR is carried out under the following cycle conditions:

Step 1: 5 minutes 94°C (denaturation)

Step 2: 3 seconds 94°C

Step 3: 1 minute 60°C (annealing)

Step 4: 2 minutes 72°C (elongation)

30 repeats of steps 2 to 4

Step 5: 10 minutes 72°C (post elongation)

Step 6: 4°C (waiting loop)

The amplicon (codARNAi-anti; SEQ ID NO: 52) is cloned using standard methods into the PCR cloning vector pGEM-T (Promega). The identity of the amplicon generated is confirmed by sequencing using the M13F (-40) primer.

Example 2 Preparation of the transgenic expression vector for expressing a codA double-stranded RNA

The codA fragments generated in example 1 are used for preparing a DNA construct suitable for expressing a double-stranded codA RNA (pSUN-codA-RNAi). The construct is suitable for reducing the steady-state RNA level of the *codA* gene in transgenic plants and, as a result therefrom, suppressing *codA* gene expression by using the double-strand RNA interference (dsRNAi) technique. For this purpose, the *codA* RNAi cassette is first constructed in the plasmid pBluKS-nitP-STLS1-35S-T and then, in a further cloning step, completely transferred to the pSUN-1 plasmid.

The vector pBluKS-nitP-STLS1-35S-T (SEQ ID NO: 55) is a derivative of pBluescript KS (Stratagene) and contains the promoter of the A. thaliana nitrilaseI gene (GenBank Acc. No.: Y07648.2, nucleotides 2456 to 4340, Hillebrand et al. (1996) Gene 170:197-200), the STLS-1 intron (Vancanneyt GF et al. (1990) Mol Gen Genet 220(2):245-250), restriction cleavage sites flanking the intron on its 5′ and 3′ sides and enabling DNA fragments to be inserted in a directed manner, and the terminator of the 35S CaMV gene (cauliflower mosaic virus; Franck et al. (1980) Cell 21:285-294). Using these restriction cleavage sites (HindIII, SalI, EcoRI, BamHI), the fragments codARNAi-sense (SEQ ID NO: 49) and codARNAi-anti (SEQ ID NO: 52) are inserted into said vector, thereby producing the finished codA RNAi cassette.

For this purpose, the codA sense fragment (codARNAi-sense SEQ ID NO: 49) is first excised from the pGEM-T vector, using the enzymes HindIII and SalI, isolated and ligated into the pBluKS-

nitP-STLS1-35S-T vector under standard conditions. This vector had previously been cleaved using the restriction enzymes HindIII and SalI. Correspondingly positive clones are identified by analytical restriction digest and sequencing.

The vector obtained (pBluKS-nitP-codAsense-STLS1-35S-T) is digested using the restriction enzymese BamHI and EcoRI. The codA-anti fragment (codARNAi-anti; SEQ ID NO: 52) is excised from the corresponding pGEM-T vector, using BamHI and EcoRI, isolated and ligated into the cut vector under standard conditions. Correspondingly positive clones which contain the complete codA-RNAi cassette (pBluKS-nitP-codAsense-STLS1-codAanti-35S-T) are identified by analytical restriction digest and sequencing.

The codA-RNAi cassette is transferred into the pSUN-1 vector (SEQ ID NO: 56) by using the SacI and KpnI restriction cleavage sites flanking the cassette. The resulting vector pSUN1-codA-RNAi (see Fig. 7; SEQ ID NO: 57) is used for transforming transgenic *A.thaliana* plants which express an active *codA* gene (see below). The plant expression vector pSUN-1 is particularly suitable within the scope of the process of the invention, since it does not contain any other positive selection marker.

The resulting vector, pSUN1-codA-RNAi, enables an artificial *codA*-dsRNA variant consisting of two identical nucleic acid elements which are separated by an intron and inverted to one another to be constitutively expressed. Transcription of this artificial *codA*-dsRNA variant results in the formation of a double-stranded RNA molecule, owing to the complementarity of the inverted nucleic acid elements. The presence of this molecule induces the suppression of *codA* gene expression (accummulation of RNA) by means of double-strand RNA interference.

Example 4: Preparation of transgenic *Arabidopis thaliana* plants

Transgenic Arabidopsis thaliana plants which express transgenically the E. coli codA gene as a marker protein ("A. thaliana-[codA]"), were prepared as described (Kirik et al. (2000) EMBO J 19(20):5562-6).

The A. thaliana-[codA] plants are transformed with an Agrobacterium tumefaciens strain (GV3101 [pMP90]) on the basis of a modified vacuum infiltration method (Clough S & Bent A (1998) Plant J 16(6):735-43; Bechtold N et al. (1993) CR Acad Sci Paris 1144(2):204-212). The *Agrobacterium tumefaciens* cells used have previously been transformed with the DNA construct described

(pSUN1-codA-RNAi). In this way, double transgenic A. thaliana-[codA] plants are generated which express an artificial codA double-stranded RNA under the control of the constitutive nitrilase1 promoter. Expression of the codA gene is suppressed as a consequence of the dsRNAi effect induced by the presence of this artificial codA-dsRNA. Said double transgenic plants may be identified owing to their regained ability to grow in the presence of 5-fluorocytosine in the culture medium.

Seeds of primary transformants are selected on the basis of the regained ability to grow in the presence of 5-fluorocytosine. For this purpose, the T1 seeds of the primary transformants are laid out on selection medium containing 200 µg/ml 5-fluorocytosine. These selection plates are incubated under long-day conditions (16 h of light, 21°C/8 h of darkness, 18°C). Seedlings which develop normally in the presence of 5-fluorocytosine are separated after 7 days and transferred to new selection plates. These plates are incubated for another 14 under unchanged conditions. The resistant seedlings are then transplanted into soil and cultured under short-day conditions (8 h of light, 21°C/16 h of darkness, 18°C). After 14 days, the young plants are transferred to the greenhouse and cultured under short-day conditions.

Example 5: Preparation of a plant transformation vector containing an expression cassette for expressing a double-stranded codA RNA and a plant selection marker

A plant selection marker consisting of a mutated variant of the A. thaliana Als gene, coding for the acetolactate synthase under the control of the promoter of the A. thaliana actin-2 gene (Meagher RB & Williamson RE (1994) The plant cytoskeleton. In The Plant Cytoskeleton (Meyerowitz, E. & Somerville, C., eds), pp. 1049-1084. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York), and the octopine synthase terminator (GIELEN J et al.(1984) EMBO J 3:835-846) is inserted into pSUN1-codA-RNAi (see Fig. 7; SEQ ID NO: 57) (At.Act.-2-At.Als-R-ocsT).

For this purpose, the pSUN1-codA-RNAi vector is first linearized using the restriction enzyme Pvu II. Subsequently, a linear DNA fragment with blunt ends, coding for a mutated variant of the acetolactate synthase (Als-R gene), is ligated into said linearized vector under standard conditions. Prior to ligation, this DNA fragment has been digested with the restriction enzyme KpnI and the protruding ends have been converted into blunt ends by treatment with Pwo DNA polymerase (Roche) according to the

manufacturer's instructions. This mutated variant of the A. thaliana Als gene cannot be inhibited by herbicides of the imidazolinone type. By expressing this mutated A.tAls-R gene, the plants obtain the ability to grow in the presence of the herbicide Pursuit™. Correspondingly positive clones (pSUN1-codA-RNAi-At.Act.-2-At.Als-R-ocsT; SEQ ID NO: 57) are identified by analytical restriction digest and sequencing.

The vector obtained enables an artificial codA RNA variant (consisting of two identical nucleic acid elements which are separated by an intron and inverted to one another) and a mutated variant of the A. thaliana Als gene to be expressed constitutively. Transcription of this artificial codA RNA variant results in the formation of a double-stranded RNA molecule, owing to the complementarity of the inverted nucleic acid elements. The presence of this molecule induces the suppression of codA gene expression (accummulation of RNA) by means of double-strand RNA interference. Expression of the Als-R gene imparts to the plants the ability to grow in the presence of herbicides of the imidazolinone type.

Example 6: Preparation of transgenic Arabidopis thaliana plants

Transgenic Arabidopsis thaliana plants expressing the E. coli codA gene as a marker protein ("A.thaliana-[codA]") were prepared as described (Kirik et al.(2000) EMBO J 19(20):5562-6).

The A.thaliana-[codA] plants are transformed with an Agrobacterium tumefaciens strain (GV3101 [pMP90]) on the basis of a modified vacuum infiltration method (Clough S & Bent A (1998) Plant J 16(6):735-43; Bechtold N et al. (1993) CR Acad Sci Paris 1144(2):204-212). The Agrobacterium tumefaciens cells used have previously been transformed with the DNA construct described (pSUN1-codA-RNAi-At.Act.-2-At.Als-R-ocsT; SEQ ID NO: 57). In this way, double transgenic A. thaliana-[codA] plants are generated which additionally express an artificial codA double-stranded RNA and a herbicide-insensitive variant of the Als gene (Als-R) under the control of the constitutive nitrilase1 promoter (A.thaliana-[codA]-[codA-RNAi-At.Act.-2-At.Als-R-ocsT]). Expression of the codA gene is suppressed as a consequence of the dsRNAi effect induced by the presence of this artificial codA-dsRNA. These double transgenic plants may be identified owing to their regained ability to grow in the presence of 5-fluorocytosine in the culture medium. In addition, positively transformed plants can be selected owing to their ability to grow in the presence of the her-

bicide Pursuit in the culture medium.

For the purpose of selection, the T1 seeds of primary transformants are therefore laid out on selection medium containing 100 µg/ml 5-fluorocytosine. These selection plates are incubated under long-day conditions (16 h of light, 21°C/8 h of darkness, 18°C). Seedlings which develop normally in the presence of 5-fluorocytosine are separated after 28 days and transferred to new selection plates. These plates are incubated for another 14 days under unchanged conditions. The resistant seedlings are then transplanted into soil and cultured under short-day conditions (8 h of light, 21°C/16 h of darkness, 18°C). After a further 14 days, the young plants are transferred to the greenhouse and cultured under short-day conditions.

In addition, seeds of the primary transformants, owing to their ability to grow in the presence of the herbicide Pursuit™, may be selected. It is furthermore possible to carry out dual selection using the herbicide Pursuit™ and 5-fluorocytosine. For this purpose, the T1 seeds of primary transformants are laid out on selection medium containing the herbicide Pursuit™ at a concentration of 100 nM (in the case of dual selection, 100 µg/ml 5-fluorocytosine is likewise present). These selection plates are incubated under long-day conditions (16 h of light, 21°C/8 h of darkness, 18°C).

Seedlings which develop normally in the presence of Pursuit™ (Pursuit™ and 5-fluorocytosine) are separated after 28 days and transferred to new selection plates. These plates are incubated under unchanged conditions for another 14 days. The resistant seedlings are then transplanted into soil and cultured under short-day conditions (8 h of light, 21°C/16 h of darkness, 18°C). After 14 days, the young plants are transferred to the greenhouse and cultured under short-day conditions.

Example 7: Analysis of the double transgenic A. thaliana plants selected using 5-fluorocytosine and/or Pursuit (A.thaliana-[codA]-[codA-RNAi- At.Act.-2-At.Als-R-ocsT])

Integration of the T-DNA region of the vector used for transformation, pSUN1-codA-RNAi-A.tAls-R, into the genomic DNA of the starting plant (A.thaliana-[codA]) and the loss of codA-specific mRNA in these transgenic plants (A.thaliana-[codA]-[codA-RNAi-At.Act.-2-At.Als-R-ocsT]) can be detected by applying Southern

analyses and PCR techniques or Northern analyses.

In order to carry out said analyses, total RNA and DNA are isolated from leaf tissue of the transgenic plants and suitable controls (using the RNeasy Maxi Kit (RNA) and Dneasy Plant Maxi Kit (genomic DNA), respectively, according to the manufacturer's information by Qiagen).

In the PCR analyses, the genomic DNA may be used directly as a basis (template) for the PCR. Total RNA is transcribed to cDNA prior to the PCR. The cDNA synthesis is carried out using the reverse transcriptase Superscript II (Invitrogen) according to the manufacturer's information.

Example 8: Detection of the reduction in the steady-state amount of codA RNA in the positively selected double transgenic plants (A.thaliana [codA]-[codA-RNAi-At.Act.-2-At.Als-R-ocsT]) in comparison with the starting plants (A.thaliana [codA]) used for transformation, by means of cDNA synthesis with subsequent PCR amplification.

PCR amplification of the codA-specific cDNA:
The cDNA of the codA gene (ACCESSION S56903) may be amplified using a sense-specific primer (codA5'C-term SEQ ID NO: 69) and an antisense-specific primer (codA3'C-term SEQ ID NO: 70). The PCR conditions to be chosen are as follows:

The PCR was carried out in 50 µl reaction mixture which contained:

- 2 µl (200 ng) of cDNA from A.thaliana -[codA] or A.thaliana [codA]-[codA-RNAi-At.Act.-2-At.Als-R-ocsT] plants
- 0.2 mM dATP, dTTP, dGTP, dCTP
- 1.5 mM $Mg(OAc)_2$
- 5 µg of bovine serum albumin
- 40 pmol of codA5'C-term SEQ ID NO: 69
- 40 pmol of codA3'C-term SEQ ID NO: 70
- 15 µl of 3.3x rTth DNA Polymerase XLPuffer (PE Applied Biosystems)
- 5U of rTth DNA Polymerase XL (PE Applied Biosystems)

The PCR was carried out under the following cycle conditions:
Step 1: 5 minutes 94°C (denaturation)
Step 2: 3 seconds 94°C
Step 3: 1 minute 56°C (annealing)

Step 4: 2 minutes 72°C (elongation)
30 repeats of steps 2 to 4
Step 5: 10 minutes 72°C (post elongation)
Step 6: 4°C (waiting loop)

In the positively selected plants, the steady-state amount of the mRNA of the codA gene and the amount of CODA protein resulting therefrom is reduced so much that a quantitative conversion of 5-fluorocytosine to 5-fluorouracil can no longer occur. Consequently, these plants (in contrast to the untransformed plants) can grow in the presence of 5-fluorocytosine. Thus it is demonstrated that transgenic plants can be identified owing to the applied principle of preventing expression of a negative selection marker.

Example 9: Detection of the DNA coding for codA-RNAi by using genomic DNA of the positively selected double transgenic plants (A.thaliana [codA]-[codA-RNAi-At.Act.-2-At.Als-R-ocsT])

The codA-RNAi transgene may be amplified using a codA-specific primer (e.g. codA5'HindIII SEQ ID NO: 50) and a 35S terminator-specific primer (35sT 5' Primer SEQ ID NO: 71). Using this primer combination, it is possible to detect specifically only the DNA coding for the codA RNAi construct, since the codA gene which was already present in the starting plants (A.thaliana [codA]) used for transformation is flanked by the nos terminator.

The PCR conditions to be chosen are as follows:
The PCR was carried out in a 50 µl reaction mixture which contains:

- 2 µl (200ng) of genomic DNA from the A.thaliana [codA]-[codA-RNAi-At.Act.-2-At.Als-R-ocsT] plants
- 0.2 mM dATP, dTTP, dGTP, dCTP
- 1.5 mM $Mg(OAc)_2$
- 5 µg of bovine serum albumin
- 40 pmol of codA-specific sense primer (SEQ ID NO: 50, 53 or 69)
- 40 pmol of 35sT 5' primer SEQ ID NO: 71
- 15 µl of 3.3x rTth DNA Polymerase XLPuffer (PE Applied Biosystems)
- 5U of rTth DNA Polymerase XL (PE Applied Biosystems)

The PCR was carried out under the following cycle conditions:
Step 1: 5 minutes 94°C (denaturation)

Step 2: 3 seconds 94°C
Step 3: 1 minute 56°C (annealing)
Step 4: 2 minutes 72°C (elongation)
30 repeats of steps 2 to 4
Step 5: 10 minutes 72°C (post elongation)
Step 6: 4°C (waiting loop)

In this way, it is possible to detect in the positively selected plants integration of the codA-RNAi DNA construct into the chromosomal DNA of the starting plants used for transformation. Thus it is demonstrated that transgenic plants can be identified owing to the applied principle of preventing expression of a negative selection marker.

Example 10: Detection of the reduction in the steady-state amount of codA RNA in the positively selected double transgenic plants (A.thaliana [codA]-[codA-RNAi-At.Act.-2-At.Als-R-ocsT]) in comparison with the starting plants (A.thaliana [codA]) used for transformation, by Northern analysis.

Gel-electrophoretic RNA fractionation:

For each RNA agarose gel, 3 g of agar are dissolved in 150 ml of H_2O (f.c. 1.5% (w/v)) in a microwave oven and cooled to 60°C. The addition of 20 ml of 10x MEN (0.2 M MOPS, 50 mM sodium acetate, 10 mM EDTA) and 30 ml of formaldehyde (f.c. 2.2 M) causes further cooling so that the well-mixed solution must be poured speedily. Formaldehyde prevents the formation of secondary structures in the RNA, and therefore the rate of migration is approximately proportional to the molecular weight (LEHRBACH H et al. (1977) Biochem J 16: 4743-4751). The RNA samples are denatured, prior to application to the gel, in the following mixture: 20 µl of RNA (1-2 µg/µl), 5 µl of 10x MEN buffer, 6 µl of formaldehyde, 20 µl of formamide.

The mixture is mixed and incubated at 65°C for 10 minutes. 1/10 volume of sample buffer and 1 µl of ethidium bromide (10 mg/ml) are added and the sample is then applied. Gel electrophoresis is carried out in horizontal gels in 1x MEN at 120 V for two to three hours. After electrophoresis, the gel is photographed under UV light with the aid of a ruler for subsequent determination of the fragment length. This is followed by blotting the RNA to a nylon membrane according to the information in: SAMBROOK J et al. Molecular cloning: A laboratory manual. Cold Spring Harbor, New York, Cold Spring Harbor Laboratory Press, 1989.

Radioactive labeling of DNA fragments and Northern hybridization

The codA cDNA fragment (codARNAi-sense SEQ ID No: 49) can be labeled using, for example, the High Prime kit sold by Roche Diagnostics. The High Prime kit is based on the "random primed" method for DNA labeling originally described by Feinberg and Vogelstein. Labeling is carried out by denaturing approx. 25 ng of DNA in 9-11 µl of H_2O at 95°C for 10 min. After a short incubation on ice, 4 µl of High Prime solution (contains a random primer mixture, 4 units of Klenow polymerase and 0.125 mM dATP, dTTP and dGTP each in a reaction buffer containing 50% glycerol) and 3-5 µl of [α32P]dCTP (30-50 µCi) are added. The reaction mixture is incubated at 37°C for at least 10 min and the unincorporated dCTP is then separated from the now radiolabeled DNA by means of gel filtration via a Sephadex G-50 column. The fragment is subsequently denatured at 95°C for 10 min and kept on ice until used. The following hybridization and preincubation buffers are used:

Hypo Hybond
250 mM sodium phosphate buffer pH 7.2
1 mM EDTA
7% SDS (g/v)
250 mM NaCl
10 µg/ml ssDNA
5% polyethylene glycol (PEG) 6000
40% formamide

The hybridization temperature when using Hypo Hybond is 42°C and the duration of hybridization is 16-24 h. The RNA filters are washed using three different solutions: 2 x SSC (300 mM NaCl; 30 mM sodium citrate) + 0.1% SDS, 1 x SSC + 0,1% SDS and 0.1 x SSC + 0.1% SDS. The duration and intensity of washing depend on the strength of the activity bond. After washing, the filters are sealed in plastic foil and an X-ray film (X-OMat, Kodak) is exposed overnight at -70°C. The signal strength on the X-ray films is a measure of the amount of codA mRNA molecules in the total RNA bound on the membranes. Thus it is possible to detect the reduction in codA mRNA in the positively selected plants compared to the starting plants used for transformation.

In the positively selected plants, the steady-state amount of the mRNA of the codA gene and the amount of CODA protein produced resulting therefrom is reduced so much that a quantitative conversion of 5-fluorocytosine to 5-fluorouracil can no longer occur. Consequently, these plants (in contrast to the untransformed plants) can grow in the presence of 5-fluorocytosine. Thus it is

demonstrated that transgenic plants can be identified owing to the applied principle of preventing expression of a negative selection marker.

Example 11: Summary of the results of "negative-negative" selection

Transformation of the codA-transgenic Arabidopsis plants with the codA-dsRNA construct (pSUN1-codA-RNAi-At.Act.-2-At.Als-R-ocsT; SEQ ID NO: 57) results in a significantly increased number of double transgenic plants into whose genome the RNAi construct has been successfully integrated, in the case of both single selection (with 5-fluorocytosine alone) and dual selection (Pursuit™ and 5-fluorocytosine) (in each case in comparison with untransformed plants). The analysis by means of PCR (see above) confirms the double transgenic state for the majority of the plants generated in this way. This successfully demonstrates the practicability of the present invention, i.e. the usability of repression of a negative marker for positive selection (more or less a "negative-negative" selection).